Evaluation of Procedures for Quality Assurance Specifications

PUBLICATION NO. FHWA-HRT-04-046 OCTOBER 2004

U.S. Department of Transportation
Federal Highway Administration

Research, Development, and Technology
Turner-Fairbank Highway Research Center
6300 Georgetown Pike
McLean, VA 22101-2296

Foreword

Much has been written to guide highway agencies in the development, implementation, and use of quality assurance specifications. Unfortunately, the guidance is scattered and piecemeal. In some cases, it is out-of-date, inconsistent, or even contradicts statistical principles. Further, agencies' negative experiences with quality assurance specifications have often not been recorded, and common mistakes are repeated by other agencies.

This report is a companion to FHWA-RD-02-095, *Optimal Procedures for Quality Assurance Specifications*. While FHWA-RD-02-095 is a manual intended to provide guidance to highway agencies, this report summarizes the research work that was performed and contains the analyses to explain and justify the provided guidance. This report will be of interest to those materials, construction, specifications, and research engineers who wish to gain a better understanding of any specific procedures recommended in the manual.

Sufficient copies of this report are being distributed to provide three copies to each FHWA Resource Center, a minimum of one copy to each FHWA Division, and a minimum of two copies to each State highway agency. Direct distribution is being made to the division offices. Additional copies for the public are available from the National Technical Information Services (NTIS), 5285 Port Royal Road, Springfield, VA 22161.

T. Paul Teng, P.E.
Director, Office of Infrastructure
Research and Development

Notice

Quality Assurance Statement

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
FHWA-HRT-04-046		

4. Title and Subtitle	5. Report Date
EVALUATION OF PROCEDURES FOR QUALITY ASSURANCE SPECIFICATIONS	October 2004
	6. Performing Organization Code
7. Author(s)	8. Performing Organization Report No.
J.L. Burati, R.M. Weed, C.S. Hughes, and H.S. Hill	
9. Performing Organization Name and Address	10. Work Unit No. (TRAIS)
Department of Civil Engineering Clemson University Clemson, SC 29634-0911	11. Contract or Grant No. DTFH61-98-C-00069
12. Sponsoring Agency's Name and Address	13. Type of Report and Period Covered
Office of Research, Development, and Technology Federal Highway Administration 6300 Georgetown Pike McLean, VA 22101-2296	Final Report Sept. 1998 – Oct. 2003
	14. Sponsoring Agency's Code

15. Supplementary Notes

Contracting Officer's Technical Representative (COTR): Peter A. Kopac, HRDI-11
This study was conducted under State Planning and Research Pooled Fund Study No. 2 (199) and was administered by the Federal Highway Administration (FHWA).

16. Abstract

The objective of this project was to develop a comprehensive quality assurance (QA) manual, supported by scientific evidence and statistical theory, which provides step-by-step procedures and instructions for developing effective and efficient QA specifications.

This technical report summarizes the steps taken to accomplish this goal, along with the analyses that were conducted to support the recommendations made in the QA manual (FHWA-RD-02-095). The analytical techniques used depended on the decision that needed to be made. Both analytical and computer simulation approaches were used.

Percent within limits (PWL) (or its complement, percent defective (PD)) was selected as the best quality measure because it combines both the sample mean and standard deviation into a single measure of quality. An approach based on a single composite quality measure derived from a general performance model to predict expected pavement life was developed and is the recommended approach for determining payment factors when multiple quality characteristics are measured. A detailed discussion and analysis are also presented regarding the risks involved in the various approaches to verifying the contractor's test results. The relatively high risks that are associated with typical agency verification testing frequencies are highlighted.

17. Key Words	18. Distribution Statement
Quality assurance, quality control, specifications, statistical specifications, QA, QC, payment adjustments.	No restrictions. This document is available to the public through the National Technical Information Service, Springfield, VA 22161.

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	414	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
		AREA		
in^2	square inches	645.2	square millimeters	mm^2
ft^2	square feet	0.093	square meters	m^2
yd^2	square yard	0.836	square meters	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	square kilometers	km^2
		VOLUME		
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	cubic meters	m^3
yd^3	cubic yards	0.765	cubic meters	m^3
		NOTE: volumes greater than 1000 L shall be shown in m^3		
		MASS		
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")
		TEMPERATURE (exact degrees)		
^oF	Fahrenheit	5 (F-32)/9 or (F-32)/1.8	Celsius	^oC
		ILLUMINATION		
fc	foot-candles	10.76	lux	lx
fl	foot-Lamberts	3.426	candela/m^2	cd/m^2
		FORCE and PRESSURE or STRESS		
lbf	poundforce	4.45	newtons	N
lbf/in^2	poundforce per square inch	6.89	kilopascals	kPa

APPROXIMATE CONVERSIONS FROM SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
		AREA		
mm^2	square millimeters	0.0016	square inches	in^2
m^2	square meters	10.764	square feet	ft^2
m^2	square meters	1.195	square yards	yd^2
ha	hectares	2.47	acres	ac
km^2	square kilometers	0.386	square miles	mi^2
		VOLUME		
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m^3	cubic meters	35.314	cubic feet	ft^3
m^3	cubic meters	1.307	cubic yards	yd^3
		MASS		
g	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg (or "t")	megagrams (or "metric ton")	1.103	short tons (2000 lb)	T
		TEMPERATURE (exact degrees)		
^oC	Celsius	1.8C+32	Fahrenheit	^oF
		ILLUMINATION		
lx	lux	0.0929	foot-candles	fc
cd/m^2	candela/m^2	0.2919	foot-Lamberts	fl
		FORCE and PRESSURE or STRESS		
N	newtons	0.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in^2

*SI is the symbol for the International System of Units. Appropriate rounding should be made to comply with Section 4 of ASTM E380. (Revised March 2003)

Preface

It is important to note that two documents have been prepared for this project—a manual for use by State highway administrations (SHAs), and this technical report, which summarizes the procedures and findings of the project. The manual is intended to be a comprehensive guide that a SHA can use when developing new or modifying existing acceptance plans and quality assurance (QA) specifications. While the focus and objectives of these documents are quite different, they are not entirely stand-alone documents. In preparing the two documents, an attempt has been made to minimize duplication of the contents. As such, this technical report should be read in conjunction with and as a companion to the QA specifications manual, *Optimal Procedures for Quality Assurance Specifications* (Report No. FHWA-RD-02-095), which also resulted from this project.

The focus of the manual is on *what* should be done when developing QA specifications. The reasons for the various steps and possible decisions are explained and easy-to-follow examples are included to assist in understanding the process that is involved. The manual does not explain what was done during the project, nor what analytical and simulation analyses were conducted, unless it was necessary to clarify why certain steps in the process were necessary. This technical report contains the detailed descriptions and summaries of the results for the analyses that were conducted to arrive at the decisions and recommendations included in the manual.

Table of Contents

Chapter **Page**

List of Figures

Figure **Page**

List of Tables

Table **Page**

1. INTRODUCTION

BACKGROUND

The majority of State highway administrations (SHAs) now employ statistical quality assurance (QA) specifications to some degree. These specifications contain statistical acceptance plans that serve to inform a contractor of at least three items: (1) the quality level that the agency desires, (2) how the contractor's submitted quality level will be determined (i.e., estimated), and (3) the consequences for the contractor when the submitted quality level estimate is below or above the desired level. Whether the acceptance plan leads to simple pass/fail decisions or adjustments in contract payment, its proper development is critical for the plan to be effective.

The development of statistical acceptance plans for highway construction requires a good understanding of statistics, materials and construction variability, and the product quality/performance/cost interrelationship. Currently, however, acceptance plan designers must make some design decisions relying more on intuition than on established engineering and economic principles. One problem, of course, is that there is much that is still unknown about highway quality, variability, performance, and cost. Some long-term research into these topics is currently underway. Nonetheless, designers must make do with whatever information, instructions, and resources are currently available when developing new, or revising existing, acceptance plans.

Although much good information exists for acceptance plan design, there is also much misinformation and confusion. This has resulted in a general lack of uniformity among highway agency acceptance plans; use of acceptance plans that range from totally ineffective to impractically severe; difficulties in evaluating the effectiveness of some nonstandard acceptance plans; and, presumably, general dissatisfaction as evidenced by frequent revisions in agency acceptance plans (these major revisions often have a significant impact on construction). It would help greatly if the acceptance plan designer had clear, specific, supported, and comprehensive guidance for developing acceptance plans based on best practices and on what is currently known. The guidance should also, where possible, take the subjectivity out of acceptance plan design and replace it with rational and defensible scientific procedures.

As part of a pooled fund study, 18 SHAs and 1 Canadian Province provided funds for this project. The agencies that provided funding for this study are shown in figure 1.

PROJECT OBJECTIVE

The objective of the project was to develop a comprehensive manual, supported by scientific evidence and statistical theory, that provides step-by-step procedures and instructions for developing effective and efficient QA specifications.

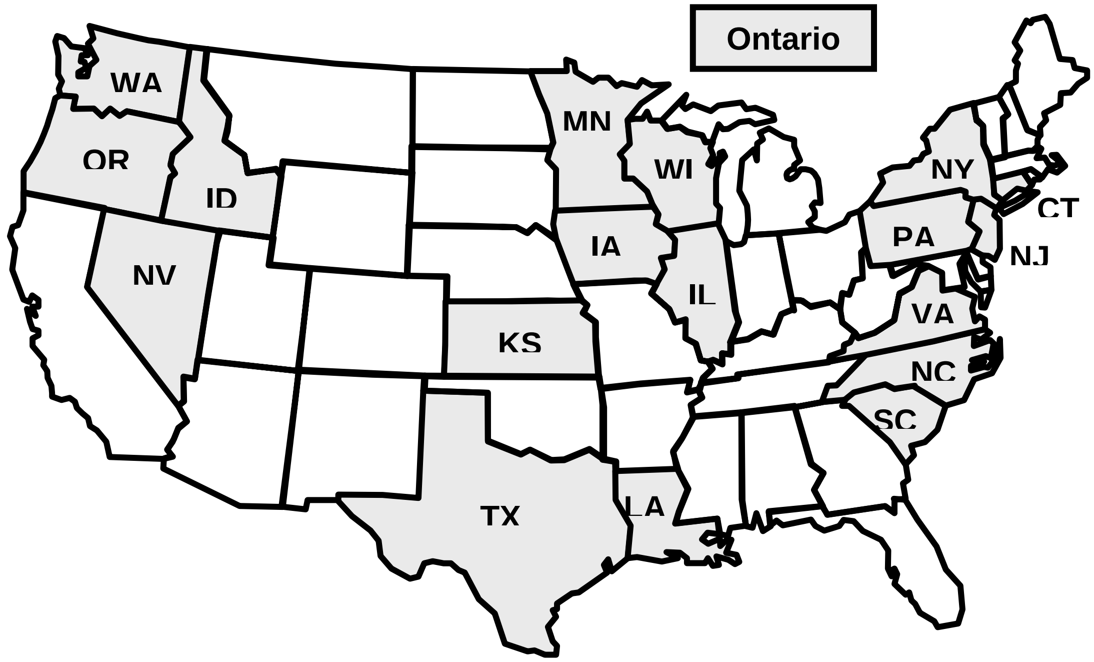


Figure 1. States that provided funding for the study.

METHODOLOGY

For the project, a team was assembled that had extensive experience in the area of QA and statistical applications to highway materials. The team included individuals with extensive experience in the development and analysis of QA and acceptance procedures. A statistician was also included to assist in the statistical analyses of the various procedures to be considered. The members of the project team are identified in table 1.

To oversee the project, a panel (subsequently referred to as *the panel*) was formed that consisted of one representative from each of the agencies that contributed funding for the study. The contracting officer's technical representative (COTR) for the Federal Highway Administration (FHWA) was also a member of the panel. The panel member for a few of the States changed over the course of the project. Table 2 shows the members comprising the panel that approved the QA specifications manual that was developed.

A number of tasks were completed in the process of accomplishing the goal of a step-by-step manual for developing QA specifications. The methodologies for accomplishing these various tasks are discussed in the following sections.

Table 1. Project team.

Member	Experience	Areas of Expertise
James L. Burati, Jr., Ph.D. Project Principal Investigator	25 years	QA Specifications, Evaluation of Asphalt Concrete (AC) Mixtures, AC Pavement Construction, Total Quality Management (TQM), Statistical Applications in Highway Construction, Computer Simulation
Hoke S. Hill, Jr., Ph.D. Project Statistical Consultant	23 years	Sampling Theory, Statistical Analysis, Statistical Graphics, Statistical Computing, Experiment Design
Richard M. Weed, P.E. Project QA Consultant	40 years	Statistical QA, Computer Simulation, Construction Specifications Development, Applications of Statistical Theory to Engineering Problems, Portland Cement Concrete (PCC) Specifications
Charles S. Hughes, P.E. Project QA Consultant	40 years	AC Mix Design, QA Specifications, Pavement Performance, Risk Assignment and Assessment, Asphalt Pavement Construction

Review of the Literature

The project team members have written a good deal of the highway materials and construction QA literature. So, the resumes of the project team were reviewed and a list of their publications relating to the proposed research was compiled. In addition, a number of computer database searches were conducted. The computer search capability of Clemson University's Cooper Library and Web-based search approaches were employed.

The abstracts obtained from the database searches were reviewed and the full publications were obtained for those abstracts that appeared to be most appropriate for further study for the current project.

In addition, all SHAs were contacted to request copies of their specifications for hot-mix asphalt concrete (HMAC) and portland cement concrete (PCC). Specifications were obtained from 31 agencies, including most of the pooled fund States for this project.

Table 2. Panel members.

Panel Member	Representing
Chris Abadie	Louisiana Department of Transportation and Development
Roger Apple	Pennsylvania Department of Transportation (DOT)
Ataur Bacchus	Ontario Ministry of Transportation (MOT)
Kevin Dayton	Washington State DOT
Steve DeWitt	North Carolina DOT
Doug Dirks	Illinois DOT
Milton Fletcher	South Carolina DOT
Steve Gage	Connecticut DOT
Jeff Hale	Nevada DOT
Kurt Johnson	Wisconsin DOT
James Klessig	Minnesota DOT
Peter Kopac	FHWA
Rick Kreider	Kansas DOT
Bill Maupin	Virginia Transportation Research Council
Garth Newman	Idaho Transportation Department
Thomas Reis	Iowa DOT
Deniz Sandhu	New York State DOT
Jeffrey Seiders	Texas DOT
Ken Stoneman	Oregon DOT
Richard Weed	New Jersey DOT

Development of the Process

One of the most important tasks for the project was to identify the necessary steps (and the specific options available at each step) for a highway agency to comprehensively develop a QA specification. To help accomplish this task, the information collected during the literature search was sorted and synthesized. The project team was also able to call on their many collective years of experience in the development and analysis of QA specifications.

The entire project team met in Clemson, SC, to develop the structure for the QA specifications development, implementation, and monitoring process. At this meeting, a detailed flowchart was developed and the major discussion points to accompany the flowchart were identified.

Subsequently, Charles Hughes developed the initial draft of an executive summary to accompany and describe the elements of the process flowchart. The executive summary and flowchart were distributed to the panel for review. The flowchart was then reviewed and discussed at a meeting attended by the panel members and members of the project team. After a great deal of discussion, some revisions were agreed upon and were subsequently incorporated into the final flowchart.

Analysis of Specific Options

To a great extent, the approved QA specifications development process flowchart identified the various analyses that needed to be conducted. A number of decisions needed to be made to progress through the flowchart. Each of these decisions had several possible options from which to select. These options were analyzed to determine which one was the best approach to recommend.

The analysis techniques used depended on the decision that needed to be made. Both analytical and computer simulation approaches were used. In some cases, such as with the risks associated with certain hypothesis tests, an analytical evaluation was relatively simple to use. In many cases, however, a computer simulation approach was the best (and, in some cases, the only) analysis method to use. This is particularly true for issues that are related to multiple quality characteristics.

The specific analyses that were conducted, and the decisions that were made based on these analyses, are covered in detail in the various chapters of this report.

Development of the Manual

The principal product of this project, as stated in the initial request for proposal (RFP), is

> *… a comprehensive manual that a highway agency can use when developing new, or modifying existing, acceptance plans. The draft manual shall provide all necessary instructions and illustrative examples to clearly lead the agency through the entire process of acceptance plan development, including:*
>
> a. *Setting up the initial data collection/experimentation to determine typical parameters of current construction.*
>
> b. *Establishing the desired level of quality to be specified.*
>
> c. *Designing the actual acceptance plan itself, including selecting the statistical quality measure, buyer's and seller's risk, lot size, number of samples, specification limits, and pay-adjustment provision.*
>
> d. *Monitoring how the acceptance plan is performing.*
>
> e. *Making necessary adjustments.*

It is important to note that two documents have been prepared for the project—a manual for SHAs and this technical report, which summarizes the procedures and findings of the project.[1] While the focus and objectives of these documents are quite different, they are not entirely stand-alone documents. In preparing the two documents, an attempt has been made to minimize duplication of coverage. *The technical report should be read in conjunction with and as a companion to the QA specifications manual that also resulted from the project.*

The QA specifications development manual is directed toward SHAs that have a need to develop new QA specifications, or revise or update existing specifications. The focus of the manual is on *what* to do when developing QA specifications. The reasons for the various steps and possible decisions are explained, and easy-to-follow examples are included to assist in understanding the process that is involved. The manual does not explain what was done during the project, nor what analytical and simulation analyses were conducted unless it was necessary to clarify why certain steps in the process are necessary.

Individual members of the project team were assigned the task of preparing the first draft of specific portions of the manual. The principal investigator then collected the individual drafts and edited them into a cohesive draft manual for review by the COTR and the panel members. A panel meeting was then held to review and discuss the final draft. Comments and suggestions resulting from the panel meeting were incorporated into the final manual.

ORGANIZATION OF THIS REPORT

Chapter 2 presents a brief overview of the literature review, with primary emphasis placed on the specifications obtained from the various agencies that responded to the request for information. Chapter 3 presents the flowchart of the QA specifications development process that guided the specific analyses that were conducted to determine and support the recommendations that are made in the QA specifications manual. The chapters that follow the flowchart present the results and conclusions from the various analyses.

2. LITERATURE AND SPECIFICATION REVIEW

LITERATURE REVIEW

One of the initial tasks associated with the project was to conduct a literature search to determine what information was currently available on the topic of QA specifications and to determine its potential relevance to the project. This task was an iterative process requiring multiple searches as the project moved forward. To encompass all of the potentially pertinent information available, a methodical approach to the literature search was adopted.

Search Methodology

Since the project team members have written extensively in the area of statistically based and QA specifications, the initial step in the literature search was to locate and review all of the information previously published by members of the research team. This yielded many published documents on such topics as sampling schemes, statistical specifications, incentive and disincentive payment schedules, acceptance procedures, statistical methods for analyzing data, and the construction of operating characteristic (OC) curves. Upon completion of this phase of the search, efforts were then directed toward outside sources of information.

The next source of information that was explored was all of the publications contained in Clemson University's Cooper Library databases. The databases were explored using various combinations of keywords, as well as subject searches. This generated numerous potential sources of information. Each possible source was reviewed and either added to the list of source documents or discarded based on the relevance of the information to the project. Upon completion of this phase, efforts were then directed toward national and international databases.

Searching these databases by keyword and subject yielded hundreds of potential titles. Among the databases searched were:

- TRIS (Transportation Research Information Service).
- Ei Compendex (engineering information).
- Dissertation Abstracts (online database of doctoral dissertations).
- ICONDA (International Construction Database).
- NTIS (National Technical Information Service).
- MathSci (American Mathematical Society).
- Engineering Materials Abstracts (engineering materials).
- Federal Research in Progress (ongoing Federal research projects).

The MathSci database was included to identify related articles on topics such as sampling plans, acceptance plans, etc., from mathematical and statistical sources such as the *Journal of Quality Technology* and the *Journal of the American Statistical Association*.

All of the titles generated from this search were examined to determine if the publications should be located and obtained. Where the titles did not provide a clear indication as to the relevance of the publication, abstracts were requested to determine if the publication was an appropriate source document.

The final phase of the literature search was to review the bibliographic contents of the collected articles to determine if publications cited in the current source documents would also be relevant to the research project. This phase yielded some additional sources.

Results

The information contained in the bibliographic sources can be separated into three major categories:

- Sampling plans.
- Price- or payment-adjustment plans.
- QA specifications and acceptance plans.

Sampling Plans: The bibliographic sources contain several different types of sampling plans. Some of the plans are currently in use for highway construction, while others were evaluated to determine whether there is an appropriate use for the plans in QA specifications. Sampling plans for which literature was reviewed included:

- Stratified random sampling.
- Random sampling.
- Zero acceptance number sampling plans.
- Variables sampling plans.
- Attributes sampling plans.
- Attributes single sampling plans.
- Attributes double sampling plans.
- Attributes multiple sampling plans.
- Attributes proportion sampling plans.
- Inverse Gaussian acceptance sampling plans.
- Conditional sampling plans.
- Bayesian acceptance sampling schemes.

Price- or Payment-Adjustment Plans: The bibliographic sources also contain a number of articles discussing price- and payment-adjustment plans. Again, some of the practices discussed are currently in use, while others were examined for potential relevance to the project. Several articles discuss adjusted payment schedules as they pertain to the quality of work and expected performance, and whether current practices are fair or can be improved. The development of price-adjustment systems is also addressed from serviceability, cost of production, value concept, and OC curve approaches. The concept of composite payment equations and their development and use is addressed in several of the bibliographic sources.

QA Specifications and Acceptance Plans: The last category of bibliographic sources is related to specifications development, acceptance procedures, and QA programs. Several articles investigate the older prescriptive highway specifications and discuss the shift toward end-result and performance-related specifications (PRS). Included is an examination of several statistical measures and how each has been used.

Summary

Not much new information was learned from the literature review. The project team members had written much of the relevant literature from the highway materials and construction industry. Most of the material from outside the highway industry was directed toward industrial applications, where more stable processes and much larger sample sizes are involved. Some of the relevant information from the literature search is discussed in subsequent chapters related to the specific analysis topics, and some is presented in the QA specifications manual that was developed for the project.

A summary of some of the documents that were reviewed during the course of the search is presented as an annotated bibliography in appendix A.

SPECIFICATION REVIEW

At the beginning of the project, all SHAs were requested to provide copies of their current specifications for HMAC and PCC pavements. In response, 23 provided copies of their specifications for HMAC, 8 provided specifications for Superpave® AC mixes, and 9 provided specifications for PCC (see table 3). Many States indicated that Superpave specifications were under development at that time. The reasons for collecting and reviewing the various agency specifications were to:

- Determine what properties were being evaluated for quality control (QC) purposes.
- Determine what properties were being evaluated for acceptance purposes.
- Determine what properties were being evaluated for payment determination.
- Determine what statistical methods were being used to determine a composite payment factor when there are multiple quality characteristics.

HMAC Specifications

A total of 23 agencies (see table 3) provided copies of their current HMAC specifications. A summary of each agency's HMAC specifications is provided in appendix B. While current at the time they were obtained, these specifications may no longer be current as of the writing of this report. Each of the agency's specifications was carefully reviewed to determine the similarities and differences in the properties evaluated for QC, acceptance, and payment determination. Additionally, the statistical measures for determining compliance and the calculations for payment factors were reviewed. This task was complicated and required many iterations because of the differing terminologies used by the various agencies. For example, some agencies use *assurance* testing for the same purpose as tests that are referred to as *acceptance* testing by other agencies.

Table 3. Agencies that provided copies of their specifications.

Agency	HMAC Specification	Superpave Specification	PCC Specification
Alaska	X		
Arkansas	X		
Colorado	X		
Connecticut		X	
Idaho	X		
Illinois	X		X
Iowa	X		X
Kansas		X	X
Louisiana		X	
Maine		X	
Maryland	X		
Michigan	X		
Minnesota	X	X	
Mississippi		X	
Montana	X		
Nebraska	X		
New Jersey			X
New York		X	
North Carolina		X	X
North Dakota	X		
Nevada	X		
Ohio	X		
Ontario	X		
Oregon	X		X
Pennsylvania	X		X
South Carolina	X		
Texas	X		X
Virginia	X		
Washington	X		
Wisconsin	X		X
Wyoming	X		
Total	**23**	**8**	**9**

Examination of the specifications revealed that the majority of the agencies use the Marshall mix design and therefore the quality characteristics evaluated for QC and acceptance are similar. The most significant difference is in the number of quality characteristics that the contractor is responsible for controlling. Some agencies require the contractor to control a few common characteristics, such as gradation, asphalt content, density, voids in the mineral aggregate (VMA), voids filled with asphalt (VFA), and total air voids. Almost all of the agencies use these characteristics for QC or acceptance. However, some agencies require control over many more characteristics, including Hveem stability, Marshall stability, Marshall flow, dust-to-asphalt ratio, maximum specific gravity (MSG), bulk specific gravity (BSG), moisture content, binder temperature, liquid limit, plastic index, fractured faces, absorption, indirect tensile strength (ITS), and tensile strength ratio (TSR). Additionally, differences in the lot sizes for testing varied widely from agency to agency. Testing frequencies are also significantly different for the various agencies. This review indicated that with the exception of a few commonly measured characteristics, the QC and acceptance procedures varied widely among the responding agencies.

The methods for determining acceptance were also investigated. The method for acceptance of material varies from agency to agency, but can be grouped into four general categories:

- Acceptance testing by the department.
- Verification of the contractor's QC tests by the department's assurance tests.
- Acceptance testing by the contractor under departmental supervision.
- Some combination of contractor and departmental testing.

The final aspects of the specifications examined were the properties evaluated and the methods for determining payment factors. Most of the responding agencies evaluate only a few properties for determining payment factors. The most common properties used are gradation, in-place density, asphalt content, VMA, and air voids. Additional properties evaluated by some agencies include Marshall stability, crushed particle count, thickness, moisture content, theoretical maximum density (TMD), laboratory-molded density, and smoothness.

Superpave Specifications

The eight States that responded with copies of their Superpave specifications are indicated in table 3. The response for this specification was low at the time because of its recent introduction as a method of mix design. A summary of each agency's Superpave specifications is provided in appendix C. The Superpave specifications were reviewed for QC, acceptance, and payment factor information.

Examination of the specifications indicated that verification of the mix design is similar for all of the agencies. Additionally, the quality characteristics evaluated for QC and acceptance do not differ substantially from agency to agency. However, there is a significant difference in the quality characteristics evaluated for acceptance. Most of the agencies evaluate the following characteristics: asphalt content, gradation, air voids, VMA, and in-place density. In addition to these commonly evaluated characteristics, three agencies evaluate mix moisture, VFA, and BSG, and two agencies evaluate TMD, dust-to-asphalt ratio, and G_{mb} @ N_{des}. At least one agency evaluates a number of other quality characteristics, such as TSR, sand equivalent, percent crushed aggregate, N_{ini}, N_{des}, and N_{max}. Four of the eight agencies use smoothness as an acceptance quality characteristic.

The methods used to determine acceptance can be grouped into two categories: acceptance testing by the agency and verification of the contractor's tests by the agency's verification tests. The most common quality characteristics used in determining payment factors are gradation, asphalt content, air voids, in-place density, and smoothness.

PCC Specifications

The nine agencies that provided copies of their PCC specifications are indicated in table 3. A summary of each agency's PCC specifications is provided in appendix D. The review of the PCC specifications revealed similarities in the quality characteristics that the contractor is responsible for controlling. Each of the agencies requires the contractor to conduct QC tests for aggregate gradation, air content, slump, unit weight/yield, and compressive or flexural strength. A majority of the agencies also require the contractor to control thickness, temperature, and smoothness. Additional characteristics evaluated by at least one agency include water-cement ratio, percent passing the 75-micrometer (μm) sieve, moisture content of the aggregate, fineness modulus, sand equivalent, fine aggregate organic impurity, and admixture dosage. The QC testing frequency requirements vary widely from agency to agency.

The quality characteristics evaluated for the acceptance of material vary only slightly from agency to agency. Most of the agencies evaluate aggregate gradation, slump, temperature, smoothness, unit weight/yield, thickness, air content, and compressive or flexural strength. The method for determining the acceptance of material varies from agency to agency, with either agency verification testing of the contractor's tests or acceptance testing by the agency. The responding agencies assigned payment factors for one or more of the following quality characteristics: smoothness, thickness, air content, compressive strength, and flexural strength.

Summary

While the review of agency specifications provided background information, not a great deal of new and useful information resulted from the review. The project team members were already familiar with many agency specifications. Also, at the panel meeting to approve the QA specification process flowchart (see next chapter), the panel decided that the QA manual would not be prescriptive with regard to which quality characteristics should be measured for QC, acceptance, or payment purposes.

3. SPECIFICATIONS DEVELOPMENT PROCESS

METHODOLOGY FOR DEVELOPING THE PROCESS

As stated in chapter 1, the objective of the project was to develop a comprehensive manual, supported by scientific evidence and statistical theory, which provides step-by-step procedures and instructions for developing effective and efficient QA specifications. Therefore, one of the most important tasks for the project was to identify the necessary steps, along with the specific options available at each step, for a SHA to comprehensively develop a QA specification.

The first step toward identifying what should be considered in the specification process was a meeting between the COTR for the project and the principal investigator. This meeting took place in Clemson, SC, and was also attended by the statistical consultant member of the project team, Hoke Hill, and the graduate research assistant who had primary responsibility for the literature search and specification review activities on the project.

The full research team then met in Clemson, SC, to develop the structure for the QA specifications development, implementation, and monitoring process. At this meeting, it was decided that the best way to present the specifications development process was in a flowchart. A detailed flowchart was developed and the major discussion points to accompany the flowchart were identified. Subsequently, one of the project team members, Charles Hughes, developed the initial draft of an executive summary to accompany and describe the elements of the process flowchart. The flowchart and the executive summary were then distributed to the members of the panel for their review.

The initial flowchart and executive summary were then reviewed and discussed at a meeting of the panel that was held in Washington, DC. After a great deal of discussion, some revisions were agreed upon. A few minor revisions to the flowchart also resulted from issues that were identified while writing the QA specifications manual.

THE PROCESS

The overall specifications development and implementation process can be divided into three primary phases:

- Phase I: Initiation and Planning.
- Phase II: Specifications Development.
- Phase III: Implementation.

The steps in each of these phases can be presented in a flowchart for each phase. The steps in each of the three phases of the overall specifications development and implementation process are presented and discussed in detail in the QA specifications manual that was developed during the project. The steps in the process are therefore not discussed in detail in this technical report.

Phase I of the specifications development and implementation process is initiation and planning. The steps that are involved in this process are identified in the flowchart in figure 2. Phase II is specifications development (the steps are identified in the flowchart in figure 3). Phase III is implementation (the steps are identified in the flowchart in figure 4).

FLOWCHARTS AS A GUIDE TO REQUIRED ANALYSES

The process flowcharts for each of the three phases of the specifications development and implementation process are presented in this technical report because the flowcharts were used to identify some of the major questions that a SHA must answer when developing a new, or modifying an existing, QA specification. Each of these questions has several possible answers. The flowcharts, therefore, helped lead to the analyses that needed to be performed on the current project to determine the recommended answers to these major questions.

There are many decisions that are required as an agency progresses through the specifications development and implementation process outlined in figures 2 through 4. Many of these decisions are subjective and depend on the specific circumstances of the SHA. Each of the steps in the process is discussed in the manual. Some of the questions, however, were best answered by further analysis of the possible options.

The following questions, identified from the specifications development and implementation process flowcharts, supported detailed analyses to be conducted during the project:

- What quality measure should be used for individual quality characteristics?
- What payment relationships should be used for individual quality characteristics?
- How should multiple quality characteristics be combined into a single payment factor?
- What procedures should be used to verify the contractor's test results if they are to be used in the acceptance and payment decisions?

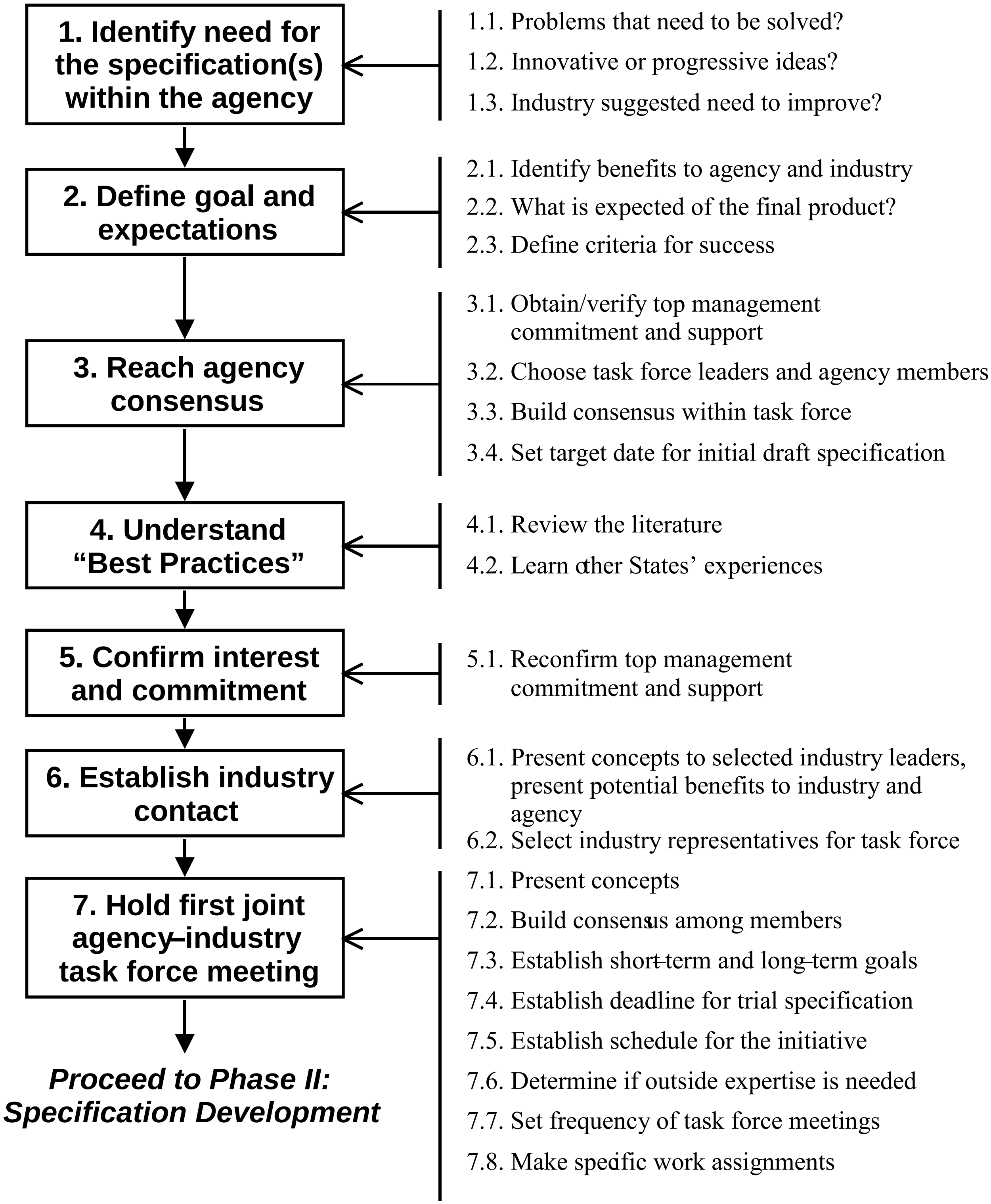


Figure 2. Flowchart for phase I—Initiation and planning.

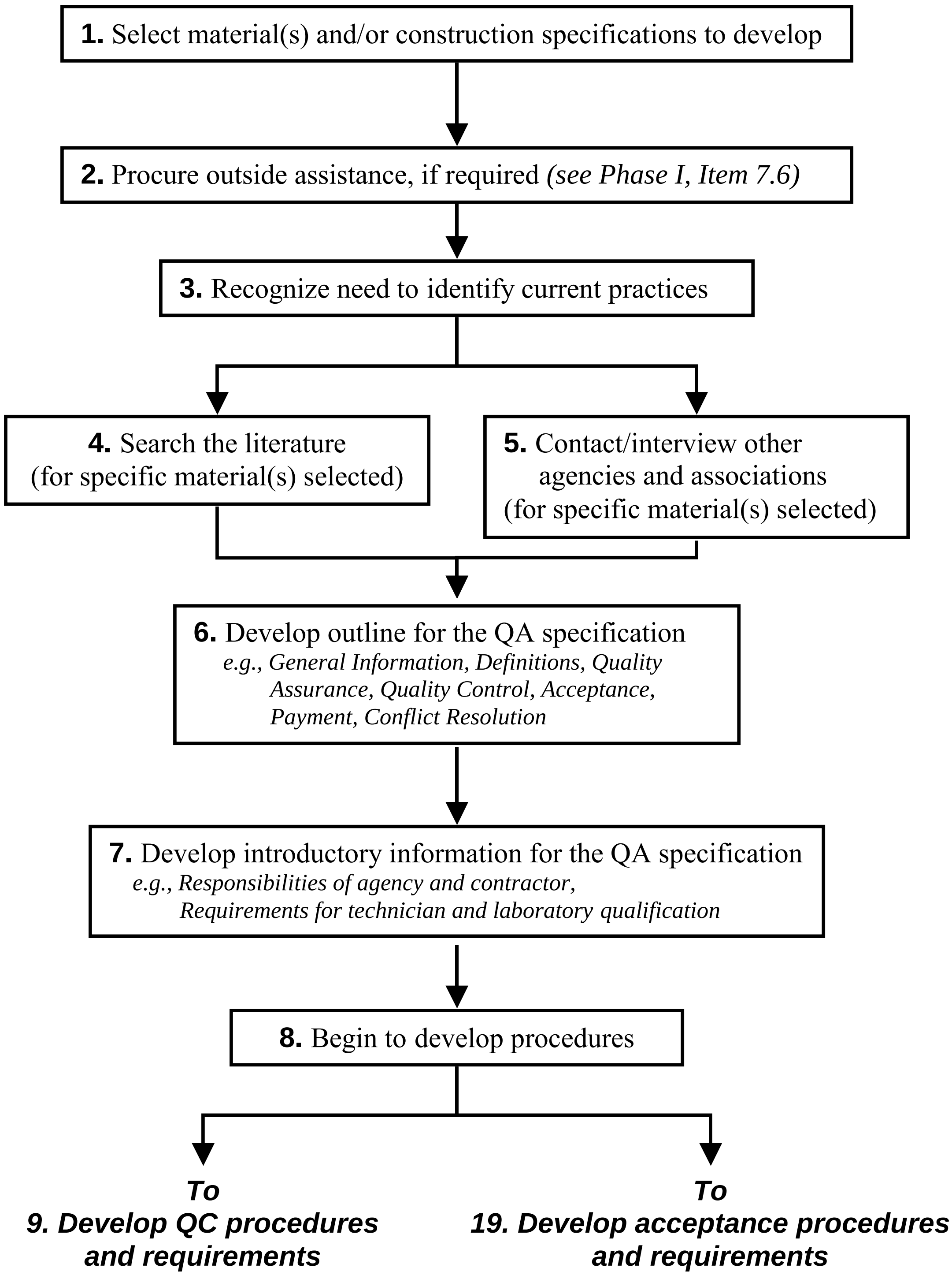


Figure 3. Flowchart for phase II—Specifications development.

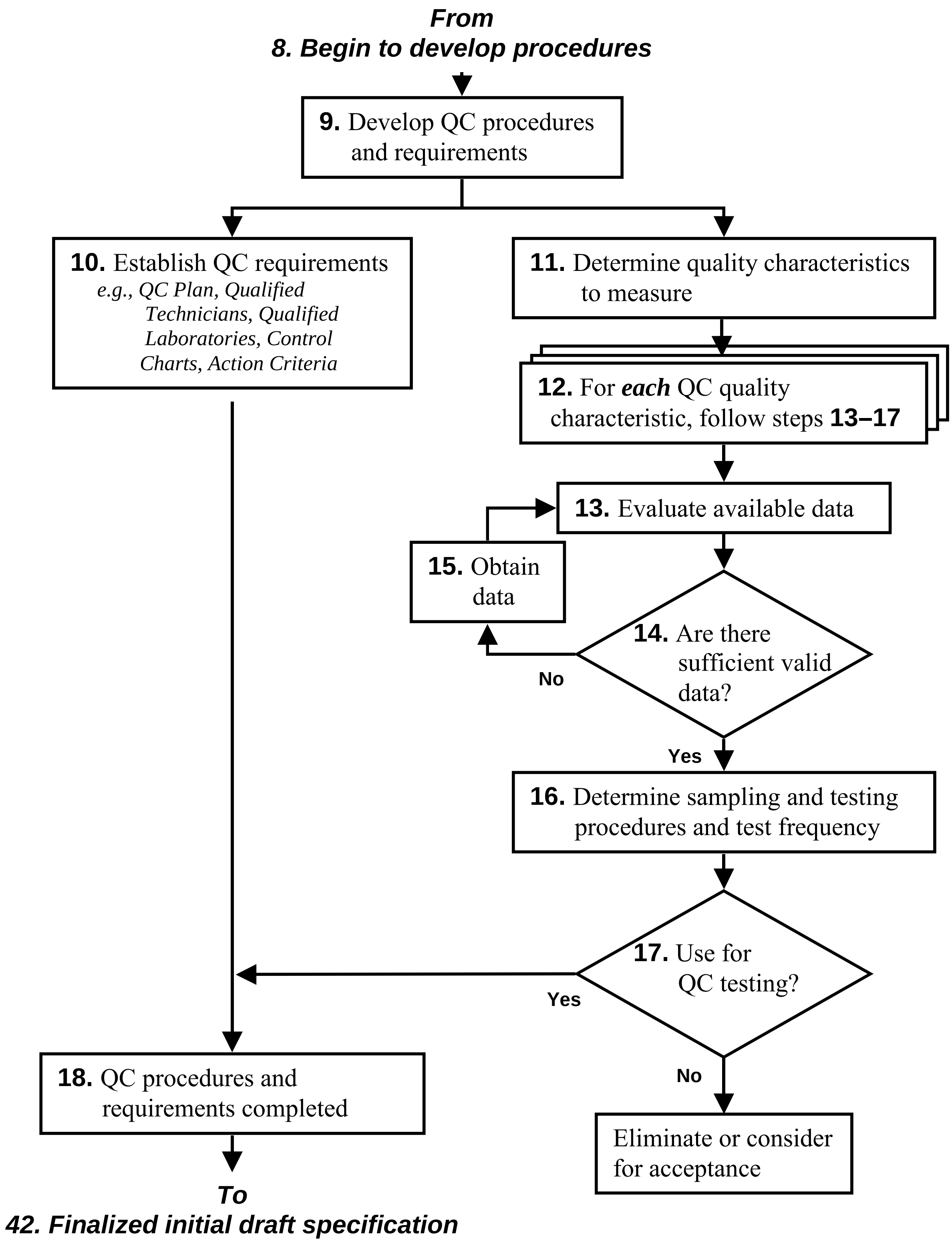


Figure 3. Flowchart for phase II—Specifications development (continued).

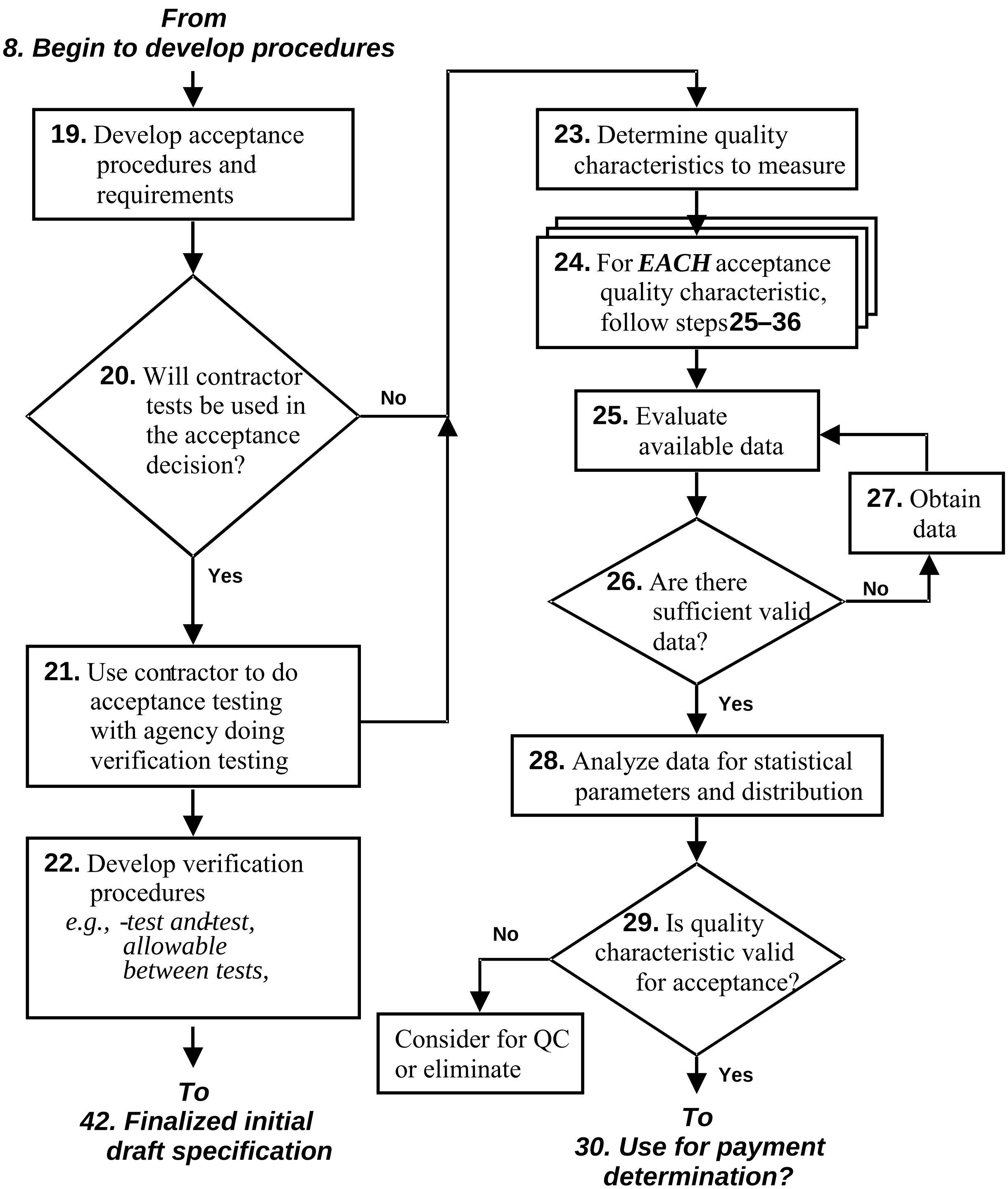


Figure 3. Flowchart for phase II—Specifications development (continued).

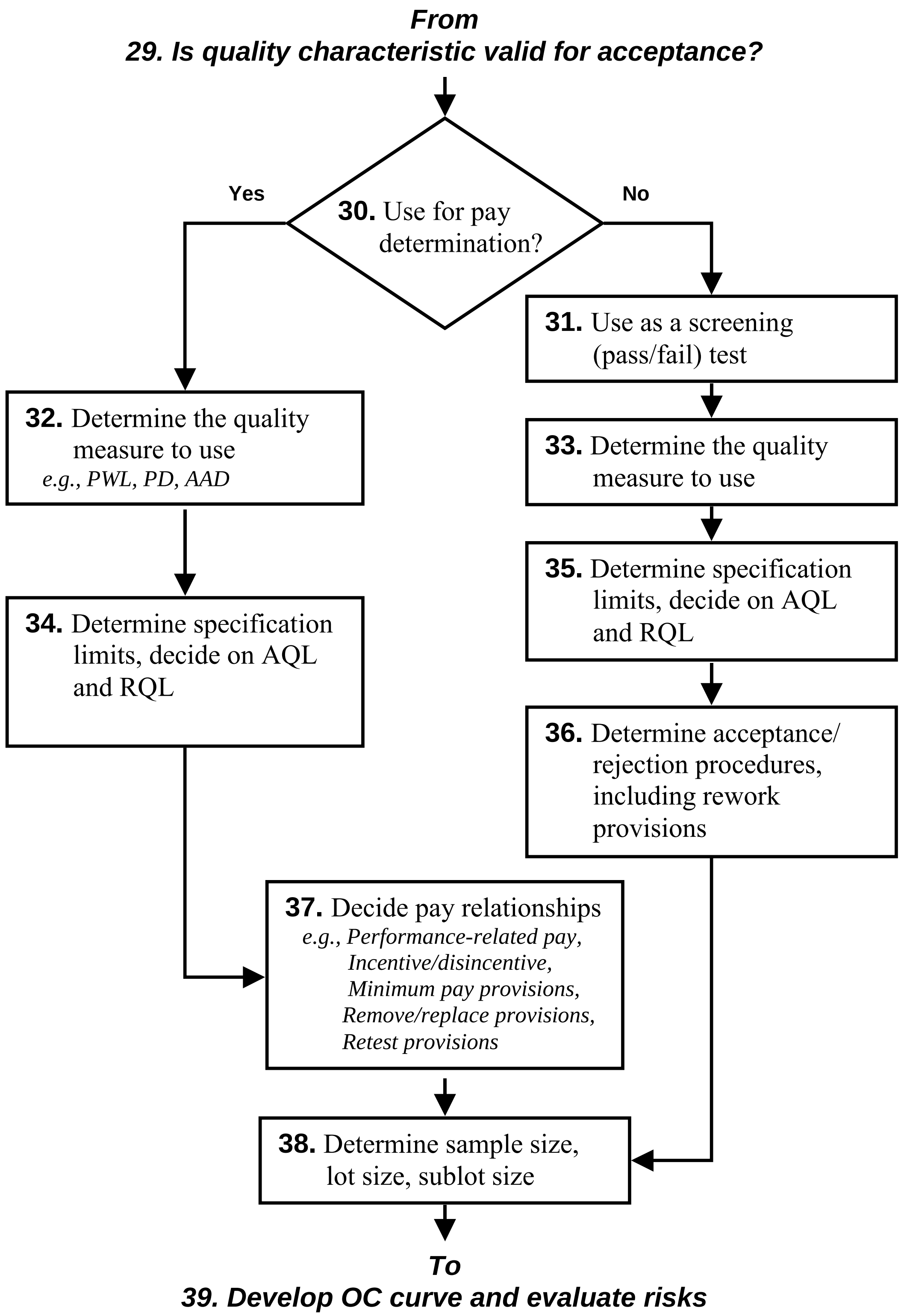


Figure 3. Flowchart for phase II—Specifications development (continued).

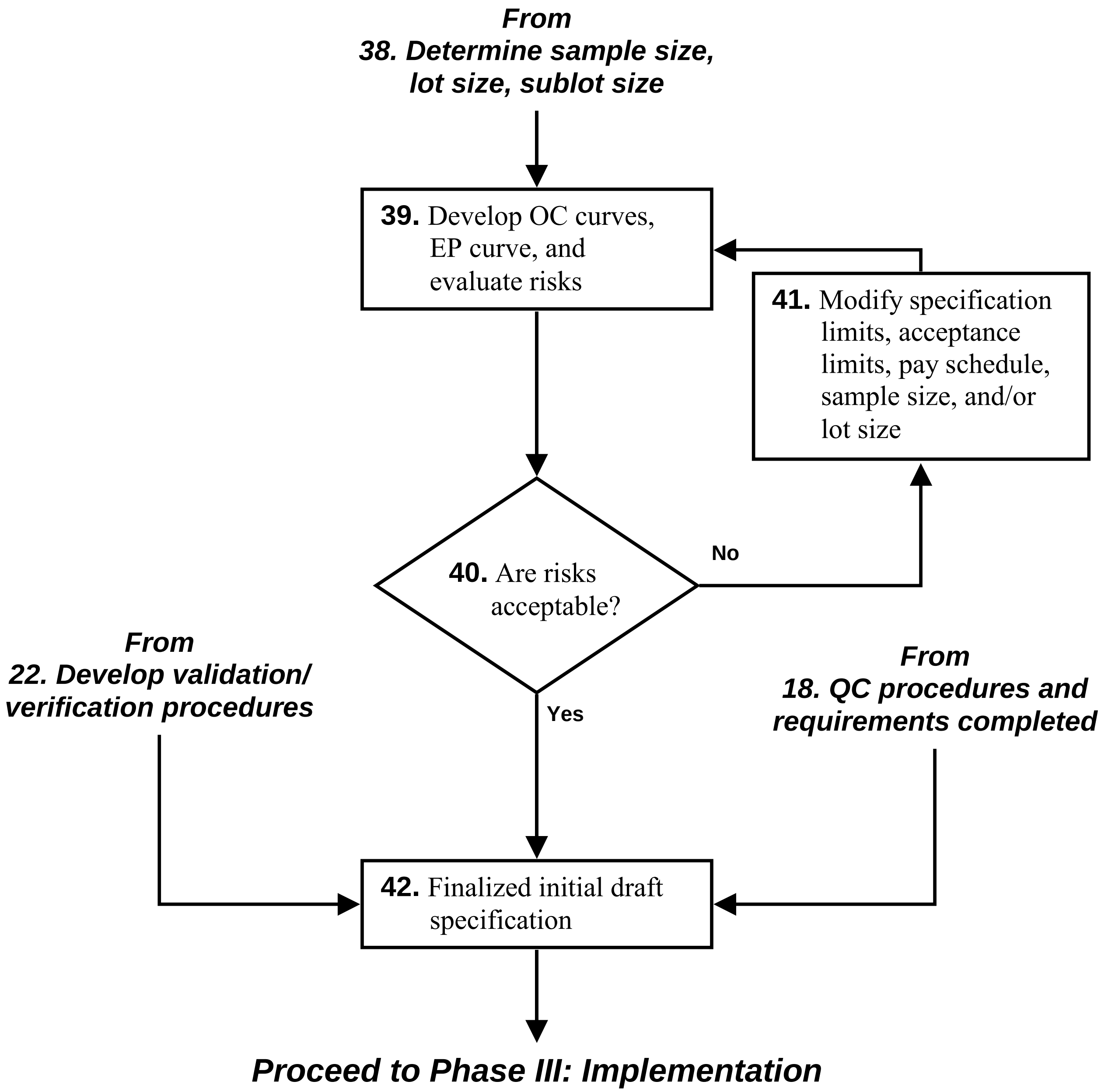


Figure 3. Flowchart for phase II—Specifications development (continued).

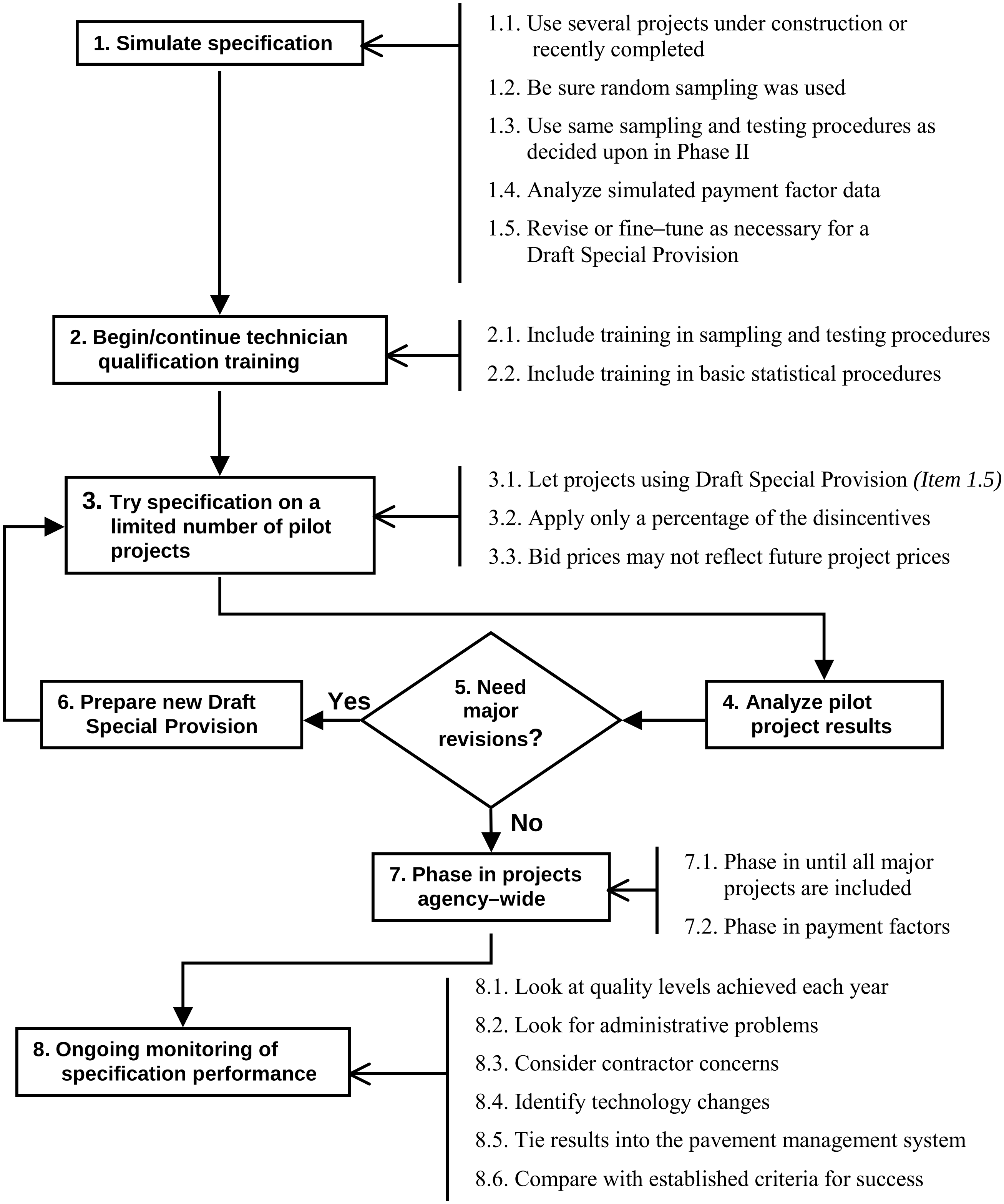


Figure 4. Flowchart for phase III—Implementation.

4. SELECTING TOPICS FOR DETAILED ANALYSES

FIRST PANEL MEETING

In accordance with contract provisions, once the specifications development flowchart was developed (see chapter 3), members of the research team met with the panel to seek its approval of the process and to identify topics for detailed analyses. The meeting took place on March 1, 1999, at the Turner-Fairbank Highway Research Center (TFHRC) in McLean, VA. The minutes of that meeting are included in appendix E.

Two major goals were planned for the meeting. The first goal was to present the preliminary specifications development flowchart to the panel members, discuss their comments and input, and obtain approval from the panel to proceed with a final process flowchart. The second goal was to determine the specific topics that the panel wanted to include for detailed analyses in the project. The minutes in appendix E indicate the process that was followed during the meeting.

With regard to the first goal, the researchers obtained input from the panel members and it was agreed that some modifications would be made to the initial flowchart. These changes were included in the final flowcharts shown in chapter 3. Concerning the second goal, there was discussion on a number of potential topics; however, there was not sufficient time for the panel to select the most desirable items for further study. It was therefore agreed that the principal investigator would distribute a survey form to the panel members to solicit their rankings of the various topics to be analyzed during the project.

SURVEY OF TOPICS FOR DETAILED ANALYSES

The principal investigator distributed a survey form to the panel members to determine a priority ranking for the various topics that were candidates for detailed analyses. The survey form that was distributed is shown in figure 5.

Of the 20 survey forms distributed (19 State representatives plus the COTR), 18 were returned. Two ranking methods were used. The first asked the respondents to group the topics into three categories—highest priority, next highest priority, and lowest priority. In summarizing these results, 5 points, 3 points, and 1 point were assigned to the topics in each category, respectively. The second ranking method asked the respondents to rank the topics in decreasing order from highest to lowest priority. In summarizing these results, 10 points were assigned to the highest priority topic, with the points decreasing to 9, 8, 7, … 3, 2, 1. Zero points were assigned to any topics that were not on the list of the top 10.

FAX TO: [contact] **FROM: ________________________**

Optimal Acceptance Procedures for Statistical Specifications

Complete the tables below using two different ranking methods. *Keep in mind that all items in the flowcharts will be addressed in the manual and the report. Some will just be addressed in general conceptual terms, while others will need to include detailed analyses to support recommendations.*

First: Rank the 4 highest priority numbered items in the table shown, along with the 4 items with second highest priority, and, finally, the 4 items with lowest priority. You may include write-in items in your priority rankings.

Priority	Numbered Items From the List
Highest (list 4)	
Next Highest (list 4)	
Lowest (list 4)	

Second: Rank the 10 highest priority numbered items in decreasing order from most important, 1, to less important, 10. You may include write-in items in your priority rankings.

Priority Ranking (1-10)	Numbered Item From the List
1 (highest)	
2	
3	
4	
5	
6	
7	
8	
9	
10 (lowest)	

Third: Cross out any of the bulleted items that you do not feel need to be included.

Fourth: Fax your ratings (and pages with crossed-out bullets) to Jim Burati at 864–656–2670.

Figure 5. Survey sent to panel members.

List of Possible Topics for Further Analysis

1) Analysis of the Percent Within Limits (PWL) approach, including:

   - Bias and precision of the PWL estimates versus sample size
   - Precision in OC curves for PWL versus sample size
   - Precision of average project PWL versus number of project lots
   - Precision of individual payments based on PWL
   - Precision of average project payment versus number of project lots
   - Effects of non-normal populations (bimodal and skewed)

2) Analysis of the Average Absolute Deviation (AAD) approach, including:

   - Bias and precision of the AAD estimates versus sample size
   - Methodology for developing and presenting AAD OC curves
   - Precision in OC curves for AAD versus sample size
   - Precision of individual payments based on AAD
   - Effects of non-normal populations (bimodal and skewed)

3) Analysis of the sample mean (mean) acceptance approach, including:

   - Bias and precision of the mean estimates versus sample size
   - Precision in OC curves for mean versus sample size
   - Precision of average project mean versus number of project lots
   - Precision of individual payments based on mean
   - Precision of average project payment versus number of project lots
   - Effects of non-normal populations (bimodal and skewed)

4) Analysis of the Conformal Index (CI) approach, including:

   - Bias and precision of the CI estimates versus sample size
   - Methodology for developing and presenting CI OC curves
   - Precision in OC curves for CI versus sample size
   - Precision of individual payments based on CI
   - Effects of non-normal populations (bimodal and skewed)

5) Analysis of the single sample variability ($\bar{X} \pm ks$) approach, including:

   - Bias and precision of the $\bar{X} \pm ks$ estimates versus sample size
   - Methodology for developing and presenting $\bar{X} \pm ks$ OC curves
   - Precision in OC curves for $\bar{X} \pm ks$ versus sample size
   - Precision of individual payments based on $\bar{X} \pm ks$
   - Effects of non-normal populations (bimodal and skewed)

Figure 5. Survey sent to panel members (continued).

6) Analysis of the moving average (m) approach for acceptance, including:

- Bias and precision of the m estimates versus sample size
- Investigation of the possibility of developing and presenting m OC curves
- Methods for applying price adjustments when using m
- Precision of individual payments based on m
- Precision of average project payment versus number of project lots
- Effects of non-normal populations (bimodal and skewed)

Note: *Some of the bulleted items for moving averages may not be possible to determine.*

7) Analysis of methods for determining lot pay factors for individual acceptance properties

8) Analysis of methods for determining composite lot pay factors when multiple acceptance properties are used

9) Analysis of the use of Bayesian procedures that incorporate information from prior lots or prior projects into the acceptance decision for the current lot

10) Analysis of procedures for verifying or validating contractor and agency test results, including:

- Use of the F-test and t-test (AASHTO QA Guide Spec.)
- Use of a single agency test and the mean and range of contractor tests (AASHTO QA Guide Spec.)
- Use of a maximum allowable difference between individual agency and contractor tests

11) Analysis of various individual "bells and whistles," that is, additional provisions that are used in conjunction with the traditional acceptance approaches, for example:

- Use of payment based on PWL but with no price reductions applied if all individual tests are within the limits
- Use of sample mean for acceptance, but also placing wider limits on individual test results
- Use of limits on sample range or standard deviation in addition to limits on the sample average
- Other provisions: ____

____

12) Other major items for analyses:

____

____

Figure 5. Survey sent to panel members (continued).

Survey Results

A summary of the survey responses is provided in table 4 for the first ranking method and in table 5 for the second ranking method. The same results are shown in graphical form, from highest to lowest priority, in figures 6 and 7 for the first and second ranking methods, respectively.

Table 4. Survey results for the first ranking method.

Agency	1	2	3	4	5	6	7	8	9	10	11	12
FHWA	1	3	1	3	1	3	5	5	5	3	1	5*
CT	3	1	3	5	1	3	5	5	1	5	3	
ID	5	1	3	3	3	1	5	5	1	5	3	
IL	5	3	5	3	3	5	1	1	1	5	3	
KS	5	5	3	5	3	5	1	1	1	3	1	3?
LA	5	1	1	1	3	3	5	5	5	3	3	
MN	5	1	1	3	5	5	3	5	1	3	3	
NV	3	5	3	1	3	3	5	5	1	5	1	
NJ	3	3	1	3	1	3	5	5	1	5	1	5+
NY	3	3	1	1	3	1	5	5	3	5	5	
ON	5	1	3	1	1	3	5	5	3	5	3	
OR	5	5	3	1	1	3	5	3	1	5	3	
PA	5	5	3	5	3	1	3	3	1	5	1	
SC	5	3	1	1	1	3	5	5	3	5	3	
TX	5	5	3	3	3	5	1	0	5	3	0	
VA	0	0	0	0	0	0	5	5	0	5	0	
WA	5	3	3	3	1	5	5	3	1	5	1	
WI	5	3	1	3	5	3	5	1	1	5	3	
Total	73	51	39	45	41	55	74	67	35	80	38	13

* Procedures for determining acceptable alpha and beta risks
? Listed an item 12 in the ranking, but did not identify it
+ Establishment of the relationship between quality/performance/value

Table 5. Survey results for the second ranking method.

Agency	1	2	3	4	5	6	7	8	9	10	11	12
FHWA	2	5	0	6	0	3	9	7	8	4	1	10*
CT	6	2	3	10	0	5	8	7	1	9	4	
ID	10	2	5	4	3	1	8	7	0	9	6	
IL	9	5	7	4	3	10	2	2	0	8	6	
KS	10	9	5	7	4	8	2	1	0	6	0	3?
LA	10	2	1	0	6	3	8	9	7	5	4	
MN	9	2	1	3	7	10	5	8	0	6	4	
NV	5	8	6	0	3	4	10	9	2	7	1	
NJ	5	3	0	4	0	6	9	8	1	7	2	10+
NY	6	5	1	2	4	0	9	10	3	7	8	
ON	10	0	3	2	1	6	8	7	5	9	4	
OR	10	7	3	2	1	4	9	6	0	8	5	
PA	10	9	5	8	4	2	6	3	1	7	0	
SC	10	4	2	0	1	3	7	8	5	9	6	
TX	9	10	5	3	4	8	2	0	7	6	0	
VA	0	0	0	0	0	0	9	9	0	9	0	
WA	7.5	5	5	5	2.5	7.5	9.5	2.5	1	9.5	0	
WI	10	5	1	6	9	4	7	2	0	8	3	
Total	138.5	83	53	66	52.5	84.5	127.5	105.5	41	133.5	54	23

* Procedures for determining acceptable alpha and beta risks
? Listed an item 12 in the ranking, but did not identify it
+ Establishment of the relationship between quality/performance/value

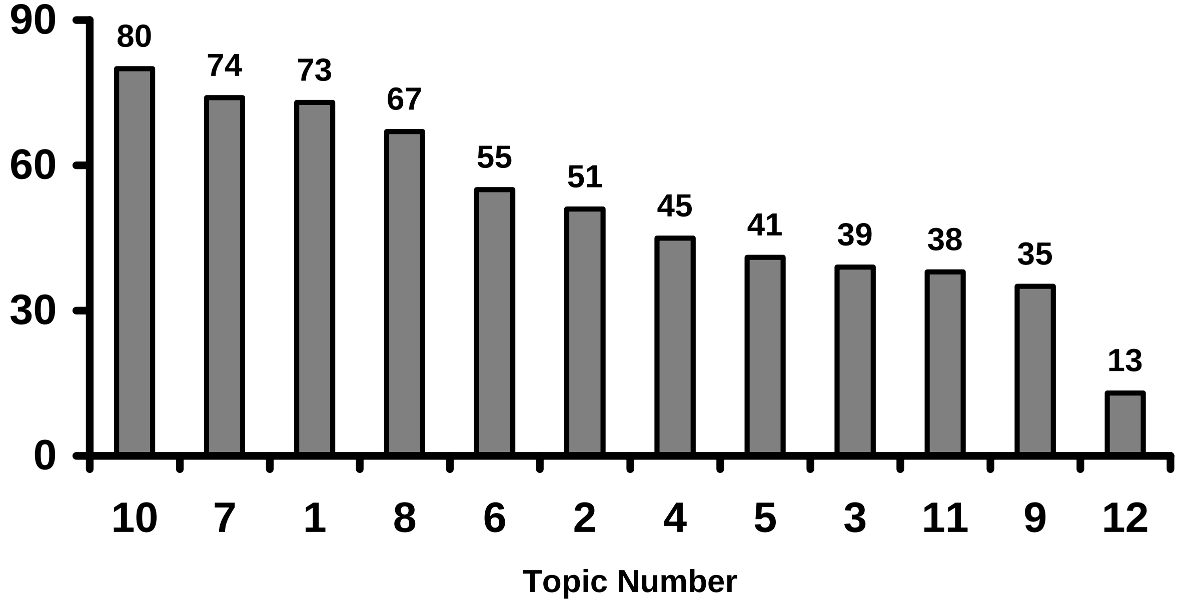


Figure 6. Graphical presentation of survey results for the first ranking method.

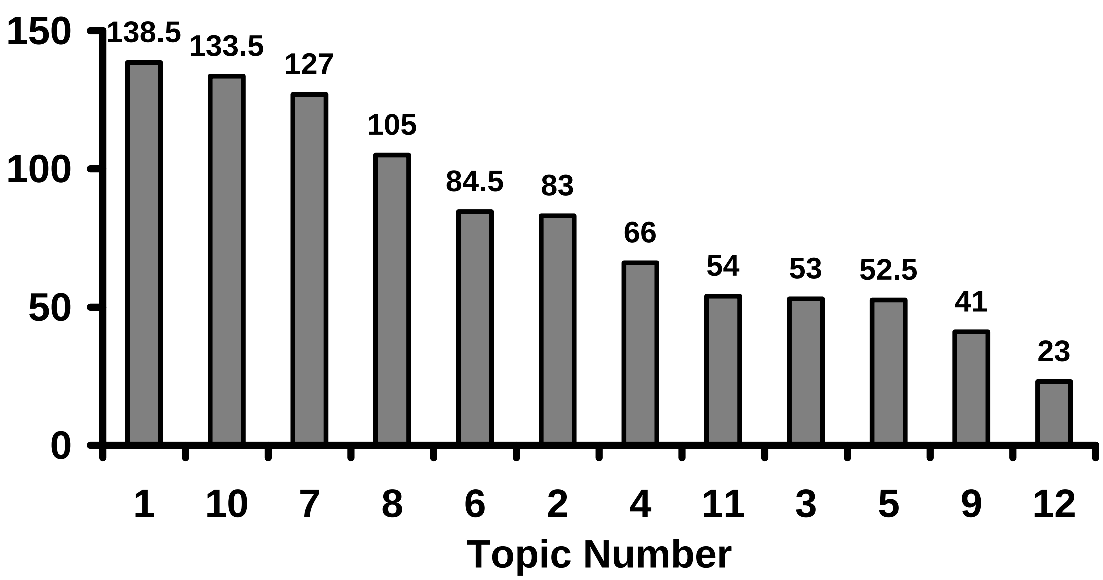


Figure 7. Graphical presentation of survey results for the second ranking method.

Table 6 shows the rankings from the two different methods and the overall ranking, which is the average of the rankings from the two ranking methods.

Table 6. Overall rankings of the survey topics.

Topic: Analysis of …		First Ranking Method	Second Ranking Method	Overall Ranking
10	Procedures for verifying or validating contractor's and agency's test results	1	2	1
1	PWL approach	3	1	2
7	Methods for determining lot pay factors for individual acceptance properties	2	3	3
8	Methods for determining composite pay factors when multiple properties are used	4	4	4
6	Moving average approach	5	5	5
2	AAD approach	6	6	6
4	CI approach	7	7	7
3	Sample mean approach	9	9	8 (tie)
5	Sample variability approach	8	10	8 (tie)
11	Various "bells and whistles"	10	8	8 (tie)
9	Use of Bayesian procedures	11	11	11

As would be expected, the two ranking methods had very similar results. The clear winners were the topics related to verifying or validating the contractor's results, the percent within limits (PWL) approach, and the determination of payment factors. There was a considerable dropoff between this group and the moving average, average absolute deviation (AAD), and conformal index (CI) approaches.

Two additional topics were proposed (each by one responder). These were "procedures for determining acceptable α and β risks" and "establishment of the relationship between quality, performance, and value." Each of these additional proposed topics would require considerable effort and, indeed, would constitute major research projects in their own right. It was not possible to tackle these topics with the time and resources that were allocated for the current project.

TOPICS SELECTED FOR DETAILED ANALYSES

Table 6 identifies the priority topics that, in the opinion of the panel, required detailed analyses during the current project. The priority items selected by the panel can be reiterated as:

- Analysis of the procedures for verifying or validating the contractor's and agency's test results.
- Analysis of the use of PWL as the quality measure.
- Analysis of the methods for determining lot pay factors for individual acceptance properties.

- Analysis of the methods for determining the composite payment factor when multiple acceptance properties are used.

These are essentially the same topics that were identified from the process flowcharts in chapter 3. Those topics were:

- What quality measure should be used for individual quality characteristics?
- What payment relationships should be used for individual quality characteristics?
- How should multiple quality characteristics be combined into a single payment factor?
- What procedures should be used to verify the contractor's test results if they are to be used in the acceptance and payment decision?

The only difference is that the panel members were interested primarily in the PWL quality measure, while the flowcharts indicate that a quality measure must be selected but do not imply that it must be PWL. Therefore, it was decided to conduct initial analyses on several potential quality measures, but to concentrate the detailed analyses on the PWL measure as long as the initial analyses indicated that it was the recommended quality measure.

Each of the bulleted items listed above is presented in depth in subsequent chapters.

5. QUALITY MEASURES: ACCURACY AND PRECISION

INTRODUCTION

There are several quality measures that can be used in an acceptance plan. In early QA specifications in the late 1960s and 1970s, the mean (or average), or the average deviation from a target value, was often used as the quality measure. However, the use of the average alone provides no measure of variability, and it is now recognized that variability is often an important predictor of performance. The problems associated with basing acceptance on only the mean are well documented. Therefore, there was no need to consider this approach in the analyses that were performed.

Several quality measures, including percent defective (PD) and PWL, have been preferred in recent years because they simultaneously measure both the average level and the variability in a statistically efficient way. Other measures that have been used by some agencies include AAD and the moving average. An additional measure that may be considered by some agencies is the CI.

PD and PWL are, in reality, the same quality measure since they are directly related by the simple relationship PWL = 100 – PD. Therefore, analyses reported using PWL also apply equally to PD, and vice versa.

The quality measures that were evaluated during the project include:

- PWL.
- AAD.
- CI.

ACCURACY AND PRECISION

Since each of the above measures provides an estimate for a true population value, it is important to know that the estimator provides an unbiased estimate for the population parameter. It is known that this is true for the procedure for estimating PWL. Computer simulation was used to evaluate whether there was bias in any of the quality measures. A lack of bias means that the estimator is accurate. Another important factor to be considered is the variability associated with the quality measure (i.e., how much do individual estimated values vary about the long-term average value for the quality measure?) This, too, was investigated by computer simulation studies. Low variability for the estimated values is a sign that the estimator is precise.

Computer simulation studies were performed to evaluate the accuracy and precision of the estimated quality measures for various sample sizes. The results of these studies are presented for each measure in the following sections.

PWL EVALUATION

The PWL for a lot can be estimated by using the quality index, *Q*. The *Q*-statistic is used with a PWL table to determine the estimated PWL for the lot. A PWL table is shown in table 7.

Conceptually, the *Q*-statistic, or quality index, performs the same function as the *Z*-statistic, except that the reference point is the mean of an individual sample, $\overline{X}$, instead of the population mean, μ, and the points of interest with regard to areas under the curve are the specification limits.

$$Q_L = \frac{\overline{X} - LSL}{s} \tag{1}$$

and

$$Q_U = \frac{USL - \overline{X}}{s} \tag{2}$$

where:
- Q_L = quality index for the lower specification limit
- Q_U = quality index for the upper specification limit
- LSL = lower specification limit
- USL = upper specification limit
- $\overline{X}$ = sample mean for the lot
- s = sample standard deviation for the lot

Q_L is used when there is a one-sided lower specification limit, while Q_U is used when there is a one-sided upper specification limit. For two-sided specification limits, the PWL value is estimated as:

$$PWL_T = PWL_U + PWL_L - 100 \tag{3}$$

where:
- PWL_U = percent below the upper specification limit (based on Q_U)
- PWL_L = percent above the lower specification limit (based on Q_L)
- PWL_T = percent within the upper and lower specification limits

Simulation Analyses

Computer simulation was used to generate samples of various sizes from known populations. These sample results were then used to estimate the PWL for the population. The mean and standard deviation for the estimated PWL values were then calculated. The difference between the mean estimated PWL was subtracted from the known population PWL to provide a measure of accuracy for the PWL estimates. The standard deviation of the estimated PWL values was used as a measure of precision (i.e., the amount of variability) for the PWL estimates.

Table 7. Quality index values for estimating PWL.

PWL	*n* = 3	*n* = 4	*n* = 5	*n* = 6	*n* = 7	*n* = 8	*n* = 9	*n* = 10 to 11
100	1.16	1.50	1.79	2.03	2.23	2.39	2.53	2.65
99	–	1.47	1.67	1.80	1.89	1.95	2.00	2.04
98	1.15	1.44	1.60	1.70	1.76	1.81	1.84	1.86
97	–	1.41	1.54	1.62	1.67	1.70	1.72	1.74
96	1.14	1.38	1.49	1.55	1.59	1.61	1.63	1.65
95	–	1.35	1.44	1.49	1.52	1.54	1.55	1.56
94	1.13	1.32	1.39	1.43	1.46	1.47	1.48	1.49
93	–	1.29	1.35	1.38	1.40	1.41	1.42	1.43
92	1.12	1.26	1.31	1.33	1.35	1.36	1.36	1.37
91	1.11	1.23	1.27	1.29	1.30	1.30	1.31	1.31
90	1.10	1.20	1.23	1.24	1.25	1.25	1.26	1.26
89	1.09	1.17	1.19	1.20	1.20	1.21	1.21	1.21
88	1.07	1.14	1.15	1.16	1.16	1.16	1.16	1.17
87	1.06	1.11	1.12	1.12	1.12	1.12	1.12	1.12
86	1.04	1.08	1.08	1.08	1.08	1.08	1.08	1.08
85	1.03	1.05	1.05	1.04	1.04	1.04	1.04	1.04
84	1.01	1.02	1.01	1.01	1.00	1.00	1.00	1.00
83	1.00	0.99	0.98	0.97	0.97	0.96	0.96	0.96
82	0.97	0.96	0.95	0.94	0.93	0.93	0.93	0.92
81	0.96	0.93	0.91	0.90	0.90	0.89	0.89	0.89
80	0.93	0.90	0.88	0.87	0.86	0.86	0.86	0.85
79	0.91	0.87	0.85	0.84	0.83	0.82	0.82	0.82
78	0.89	0.84	0.82	0.80	0.80	0.79	0.79	0.79
77	0.87	0.81	0.78	0.77	0.76	0.76	0.76	0.75
76	0.84	0.78	0.75	0.74	0.73	0.73	0.72	0.72
75	0.82	0.75	0.72	0.71	0.70	0.70	0.69	0.69
74	0.79	0.72	0.69	0.68	0.67	0.66	0.66	0.66
73	0.76	0.69	0.66	0.65	0.64	0.63	0.63	0.63
72	0.74	0.66	0.63	0.62	0.61	0.60	0.60	0.60
71	0.71	0.63	0.60	0.59	0.58	0.57	0.57	0.57
70	0.68	0.60	0.57	0.56	0.55	0.55	0.54	0.54
69	0.65	0.57	0.54	0.53	0.52	0.52	0.51	0.51
68	0.62	0.54	0.51	0.50	0.49	0.49	0.48	0.48
67	0.59	0.51	0.47	0.47	0.46	0.46	0.46	0.45
66	0.56	0.48	0.45	0.44	0.44	0.43	0.43	0.43
65	0.52	0.45	0.43	0.41	0.41	0.40	0.40	0.40
64	0.49	0.42	0.40	0.39	0.38	0.38	0.37	0.37
63	0.46	0.39	0.37	0.36	0.35	0.35	0.35	0.34
62	0.43	0.36	0.34	0.33	0.32	0.32	0.32	0.32
61	0.39	0.33	0.31	0.30	0.30	0.29	0.29	0.29
60	0.36	0.30	0.28	0.27	0.27	0.27	0.26	0.26
59	0.32	0.27	0.25	0.25	0.24	0.24	0.24	0.24
58	0.29	0.24	0.23	0.22	0.21	0.21	0.21	0.21
57	0.25	0.21	0.20	0.19	0.19	0.19	0.18	0.18
56	0.22	0.18	0.17	0.16	0.16	0.16	0.16	0.16
55	0.18	0.15	0.14	0.14	0.13	0.13	0.13	0.13
54	0.14	0.12	0.11	0.11	0.11	0.11	0.10	0.10
53	0.11	0.09	0.08	0.08	0.08	0.08	0.08	0.08
52	0.07	0.06	0.06	0.05	0.05	0.05	0.05	0.05
51	0.04	0.03	0.03	0.03	0.03	0.03	0.03	0.03
50	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 7. Quality index values for estimating PWL (continued).

PWL	*n* = 12 to 14	*n* = 15 to 18	*n* = 19 to 25	*n* = 26 to 37	*n* = 38 to 69	*n* = 70 to 200	*n* = 201 to ∞
100	2.83	3.03	3.20	3.38	3.54	3.70	3.83
99	2.09	2.14	2.18	2.22	2.26	2.29	2.31
98	1.91	1.93	1.96	1.99	2.01	2.03	2.05
97	1.77	1.79	1.81	1.83	1.85	1.86	1.87
96	1.67	1.68	1.70	1.71	1.73	1.74	1.75
95	1.58	1.59	1.61	1.62	1.63	1.63	1.64
94	1.50	1.51	1.52	1.53	1.54	1.55	1.55
93	1.44	1.44	1.45	1.46	1.46	1.47	1.47
92	1.37	1.38	1.39	1.39	1.40	1.40	1.40
91	1.32	1.32	1.33	1.33	1.33	1.34	1.34
90	1.26	1.27	1.27	1.27	1.28	1.28	1.28
89	1.21	1.22	1.22	1.22	1.22	1.22	1.23
88	1.17	1.17	1.17	1.17	1.17	1.17	1.17
87	1.12	1.12	1.12	1.12	1.12	1.13	1.13
86	1.08	1.08	1.08	1.08	1.08	1.08	1.08
85	1.04	1.04	1.04	1.04	1.04	1.04	1.04
84	1.00	1.00	1.00	1.00	0.99	0.99	0.99
83	0.96	0.96	0.96	0.96	0.95	0.95	0.95
82	0.92	0.92	0.92	0.92	0.92	0.92	0.92
81	0.89	0.88	0.88	0.88	0.88	0.88	0.88
80	0.85	0.85	0.85	0.84	0.84	0.84	0.84
79	0.82	0.81	0.81	0.81	0.81	0.81	0.81
78	0.78	0.78	0.78	0.78	0.77	0.77	0.77
77	0.75	0.75	0.75	0.74	0.74	0.74	0.74
76	0.72	0.71	0.71	0.71	0.71	0.71	0.71
75	0.69	0.68	0.68	0.68	0.68	0.68	0.67
74	0.66	0.65	0.65	0.65	0.65	0.64	0.64
73	0.62	0.62	0.62	0.62	0.62	0.61	0.61
72	0.59	0.59	0.59	0.59	0.59	0.58	0.58
71	0.57	0.56	0.56	0.56	0.56	0.55	0.55
70	0.54	0.53	0.53	0.53	0.53	0.53	0.52
69	0.51	0.50	0.50	0.50	0.50	0.50	0.50
68	0.48	0.48	0.47	0.47	0.47	0.47	0.47
67	0.45	0.45	0.45	0.44	0.44	0.44	0.44
66	0.42	0.42	0.42	0.42	0.41	0.41	0.41
65	0.40	0.39	0.39	0.39	0.39	0.39	0.39
64	0.37	0.36	0.36	0.36	0.36	0.36	0.36
63	0.34	0.34	0.34	0.34	0.33	0.33	0.33
62	0.31	0.31	0.31	0.31	0.31	0.31	0.31
61	0.29	0.29	0.28	0.28	0.28	0.28	0.28
60	0.26	0.26	0.26	0.26	0.26	0.25	0.25
59	0.23	0.23	0.23	0.23	0.23	0.23	0.23
58	0.21	0.21	0.20	0.20	0.20	0.20	0.20
57	0.18	0.18	0.18	0.18	0.18	0.18	0.18
56	0.16	0.15	0.15	0.15	0.15	0.15	0.15
55	0.13	0.13	0.13	0.13	0.13	0.13	0.13
54	0.10	0.10	0.10	0.10	0.10	0.10	0.10
53	0.08	0.08	0.08	0.08	0.08	0.08	0.08
52	0.05	0.05	0.05	0.05	0.05	0.05	0.05
51	0.03	0.03	0.03	0.03	0.03	0.03	0.02
50	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Source: "Specification Conformity Analysis," *FHWA Technical Advisory T5080.12*, June 23, 1989

As was known in advance, the quality index provided an unbiased estimate of the population mean, as evidenced by the lack of bias in the mean of the estimated PWL values. Also as expected, the amount of variability in the PWL estimates decreased as the sample size increased. This is shown in table 8. Also shown in the table is the fact that the standard deviation of the estimated PWL values is a maximum at an actual PWL of 50, and decreases to a minimum as the actual PWL approaches 100 or zero. This is also shown in the plot in figure 8.

The reason for this is illustrated in figures 9 through 11, which plot histograms of the distribution of estimated PWL values for sample sizes = 3, 5, and 10, and for actual PWL values = 90, 70, and 50. As shown in the figures, as the actual PWL approaches the natural boundary imposed by 100 PWL, the spread of estimated values decreases since the estimated values cannot exceed 100 (see figure 9). As the actual PWL value moves away from the natural boundary of 100 to a value of 70, there is a greater spread available for the estimated values (see figure 10). When the actual PWL value is 50, there is the maximum opportunity for the estimated values to spread out before hitting the natural boundaries of 100 and 0 PWL (see figure 11).

To further illustrate the lack of bias and how the variability of PWL estimates varies with sample size, figure 12 plots the average estimated PWL values from the simulation results versus the actual PWL values. The straight line at a 45-degree angle, which indicates that the estimated values vary only very slightly from the actual PWL values, shows the lack of bias in the estimator. The other lines on each plot represent the 10^{th} and 90^{th} percentiles (i.e., the values that 10 percent of the estimates are either below or exceed, respectively). As the sample size increases, these bounds become narrower, indicating a reduction in variability that is also reflected in the standard deviation values in table 7. Also, figure 13 plots the average differences versus the actual PWL value to show that the differences are centered on zero and show no consistent positive or negative bias.

To investigate the effect of project size, for sample sizes = 3, 5, and 10, various numbers of lots per project, ranging from 1 to 90, were also considered. As would be expected, since it represented an increase in the total number of tests, the amount of variability in the average estimated PWL decreased as the number of lots in the project increased. This is shown in tables 9 through 11 and in figure 14.

Table 8. Accuracy and precision for PWL estimates (based on the results of 10,000 simulated lots).

Tests per Lot, *n*	Actual PWL of Population	Mean of Estimated PWL Values Minus Actual PWL	Standard Deviation of Estimated PWL Values
3	95	–0.08	10.93
	90	+0.29	14.78
	85	–0.16	18.48
	80	–0.21	20.89
	70	–0.21	23.51
	60	+0.28	25.31
	50	–0.26	26.52
	40	–0.60	25.50
	30	–0.05	24.23
	20	+0.24	20.67
	15	–0.01	18.13
	10	–0.30	15.31
	5	+0.10	10.79
5	95	–0.12	7.94
	90	–0.43	11.42
	85	–0.26	13.42
	80	+0.08	15.15
	70	–0.18	17.10
	60	+0.13	17.71
	50	+0.24	18.39
	40	–0.43	17.89
	30	+0.08	17.71
	20	+0.10	14.96
	15	–0.32	13.49
	10	+0.30	11.84
	5	+0.18	7.88
10	95	–0.18	5.45
	90	+0.25	7.08
	85	+0.24	9.28
	80	–0.12	10.38
	70	+0.08	11.85
	60	+0.07	12.50
	50	–0.30	13.01
	40	+0.27	12.47
	30	–0.26	11.89
	20	–0.06	10.65
	15	–0.24	9.02
	10	+0.23	7.88
	5	–0.27	4.89

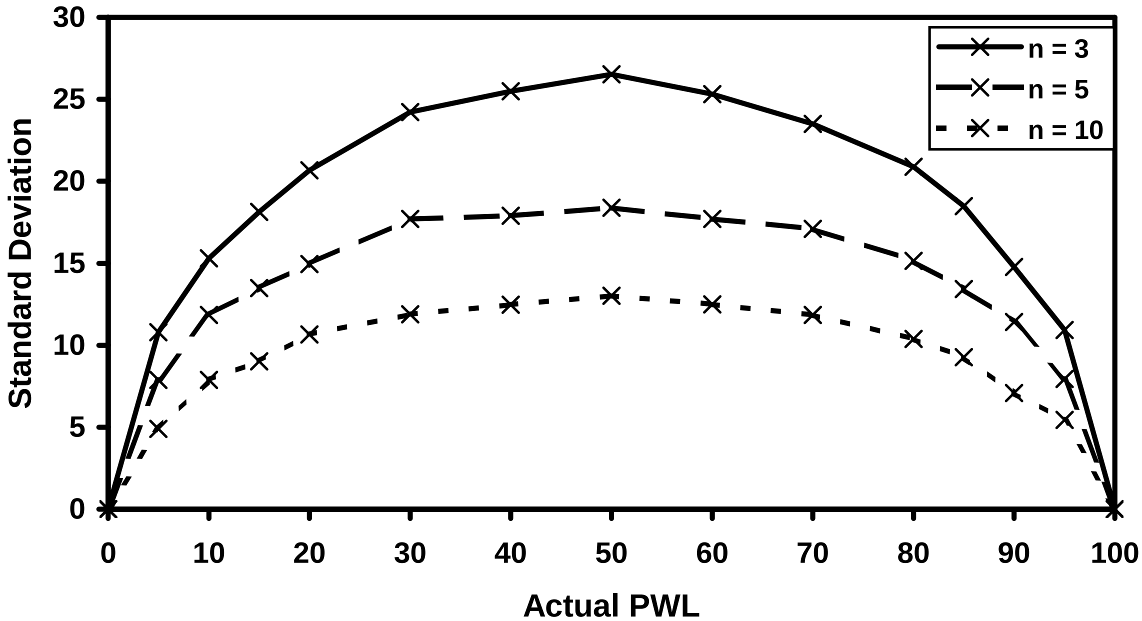


Figure 8. Plot of how the standard deviation of PWL estimates varies with the population PWL.

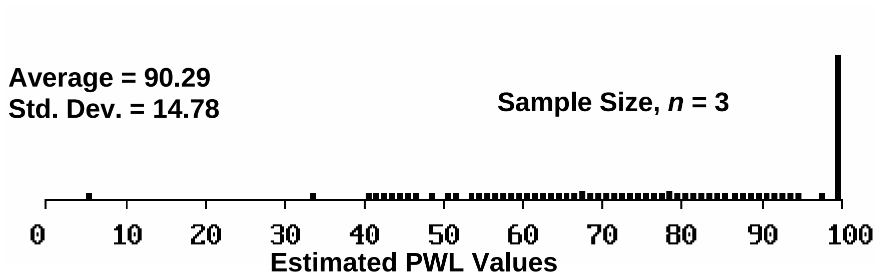


Figure 9a. Histogram illustrating the distribution of estimated PWL values for 1000 simulated lots from a population with 90 PWL, sample 3.

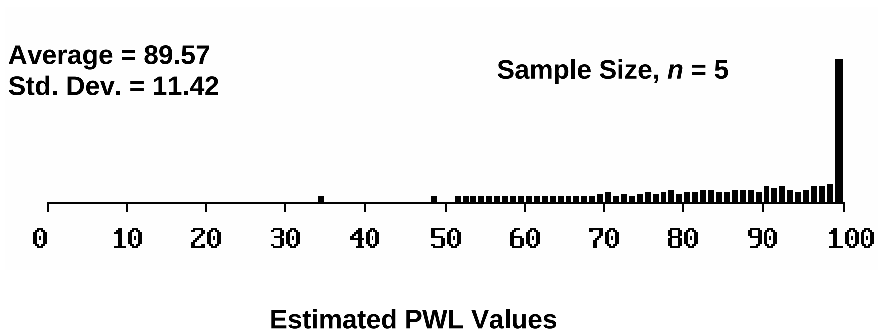


Figure 9b. Histogram illustrating the distribution of estimated PWL values for 1000 simulated lots from a population with 90 PWL, sample 5.

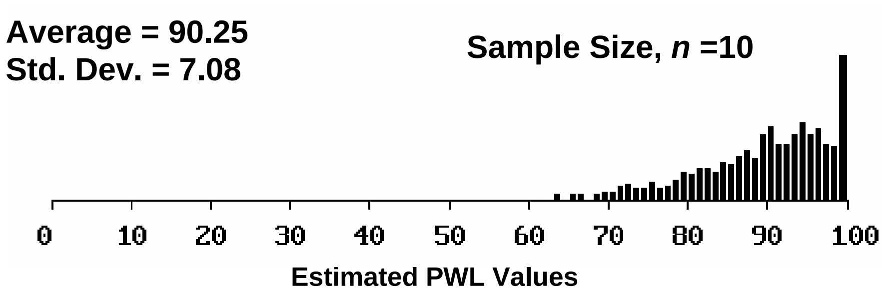


Figure 9c. Histogram illustrating the distribution of estimated PWL values for 1000 simulated lots from a population with 90 PWL, sample 10.

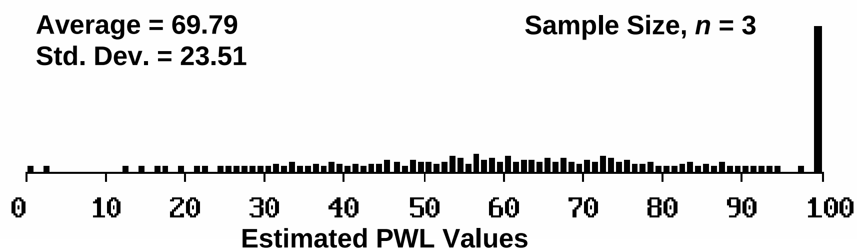


Figure 10a. Histogram illustrating the distribution of estimated PWL values for 1000 simulated lots from a population with 70 PWL, sample 3.

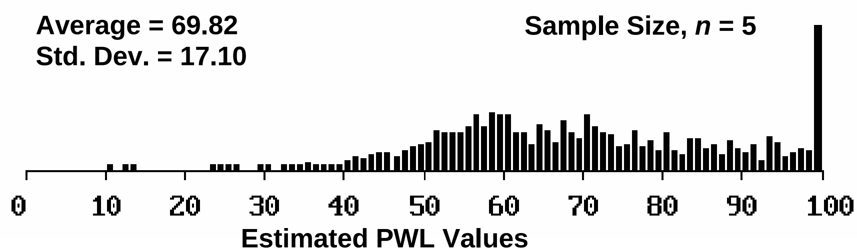


Figure 10b. Histogram illustrating the distribution of estimated PWL values for 1000 simulated lots from a population with 70 PWL, sample 5.

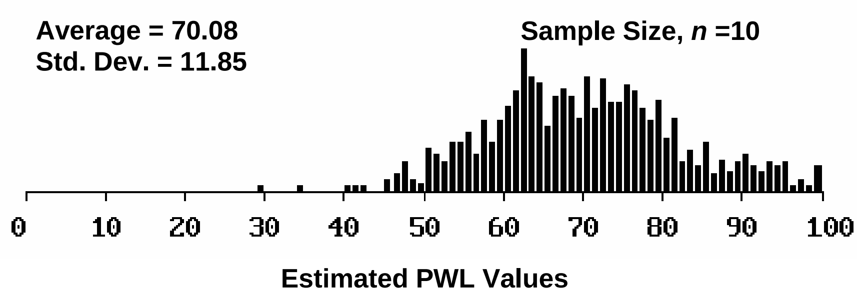


Figure 10c. Histogram illustrating the distribution of estimated PWL values for 1000 simulated lots from a population with 70 PWL, sample 10.

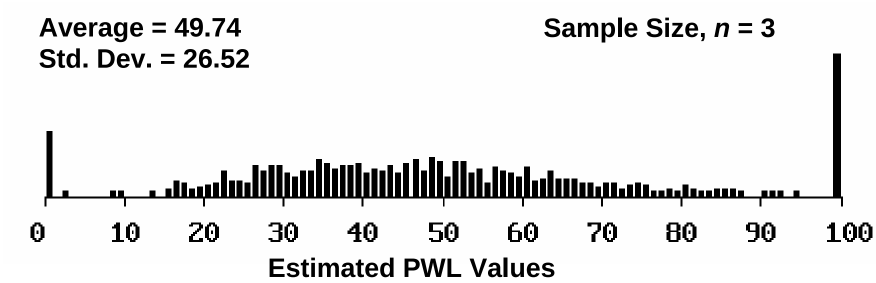


Figure 11a. Histogram illustrating the distribution of estimated PWL values for 1000 simulated lots from a population with 50 PWL, sample 3.

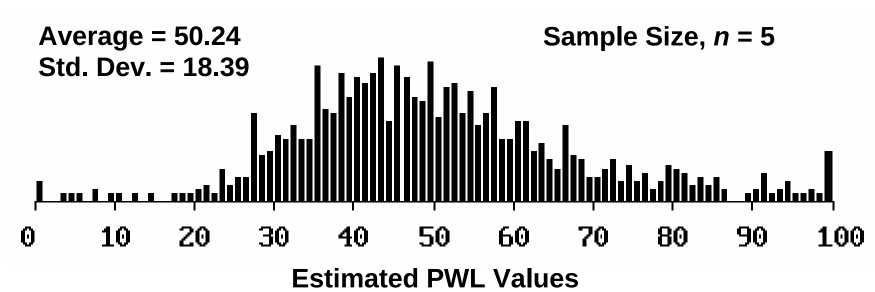


Figure 11b. Histogram illustrating the distribution of estimated PWL values for 1000 simulated lots from a population with 50 PWL, sample 5.

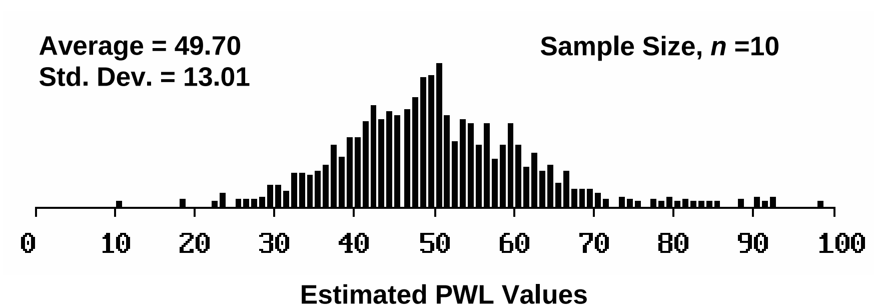


Figure 11c. Histogram illustrating the distribution of estimated PWL values for 1000 simulated lots from a population with 50 PWL, sample 10.

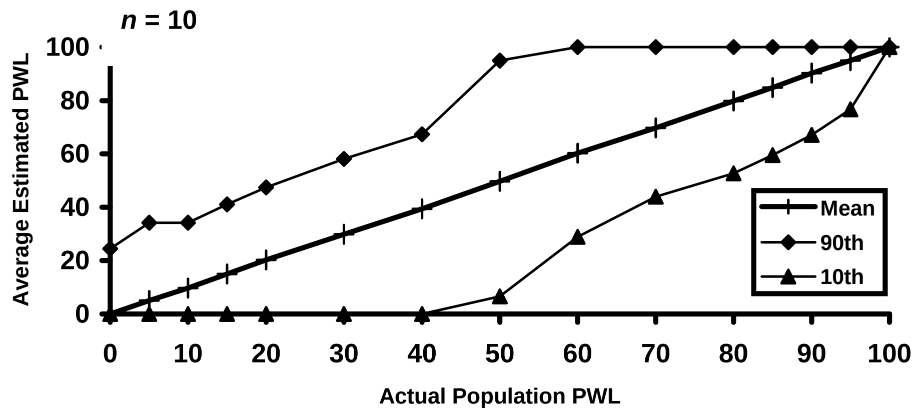


Figure 12a. Illustration 1 of accuracy and precision of PWL estimates.

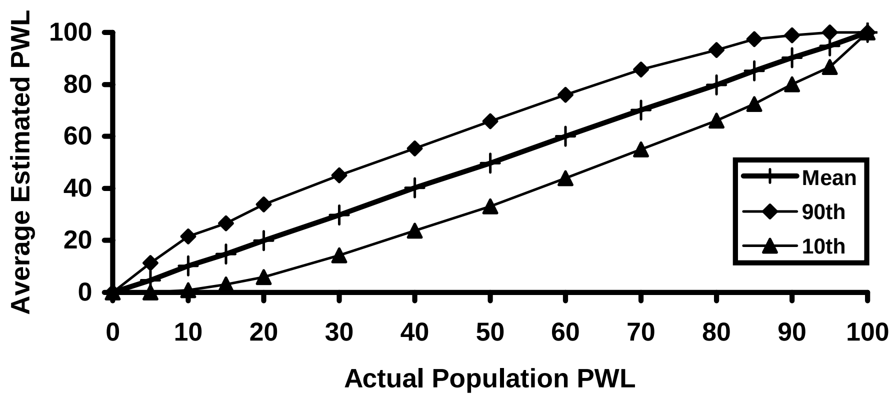


Figure 12b. Illustration 2 of accuracy and precision of PWL estimates.

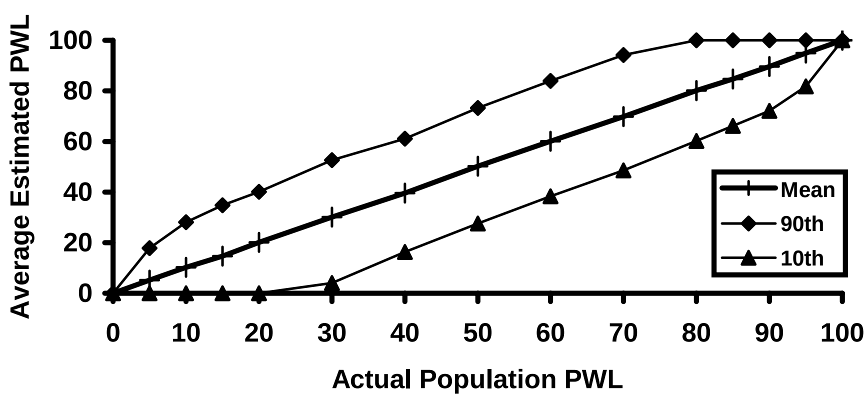


Figure 12c. Illustration 3 of accuracy and precision of PWL estimates.

Table 9. Results of simulation analyses with actual PWL = 90 (distribution of the results from 1000 simulated projects).

Tests/Lot	Lots/Project	Tests/Project	Avg. Diff.	2.5%	97.5%
3	1	3	+0.1	–10.0	+35.0
	3	9	+0.1	–10.0	+20.3
	5	15	–0.3	–10.0	+14.3
	10	30	–0.2	–9.0	+9.9
	20	60	+0.1	–6.5	+7.6
	30	90	–0.0	–5.1	+5.9
	40	120	–0.1	–4.7	+4.8
	50	150	+0.0	–4.1	+4.6
	60	180	–0.1	–3.6	+3.9
	70	210	+0.0	–3.4	+3.7
	80	240	–0.1	–3.4	+3.4
	90	270	–0.0	–2.9	+3.1
5	1	5	–0.2	–10.0	+25.5
	3	15	–0.0	–10.0	+15.5
	5	25	+0.0	–8.3	+10.5
	10	50	+0.2	–6.2	+7.5
	20	100	+0.0	–4.6	+5.6
	30	150	+0.1	–3.8	+4.2
	40	200	–0.1	–3.4	+3.7
	50	250	–0.0	–3.2	+3.1
	60	300	+0.0	–2.8	+2.8
10	1	10	+0.1	–10.0	+16.6
	3	30	+0.2	–7.6	+9.8
	5	50	+0.0	–5.7	+7.2
	10	100	–0.1	–4.4	+4.7
	20	200	–0.1	–3.3	+3.1
	30	300	+0.1	–2.7	+2.9
	40	400	–0.1	–2.4	+2.4
	50	500	–0.0	–2.2	+2.1

Avg. Diff. = average difference between the average estimated PWL and the actual PWL for each simulated project

2.5% = value for which 2.5 percent of the simulated average differences were less than or equal

97.5% = value for which 97.5 percent of the simulated average differences were less than or equal

Table 10. Results of simulation analyses with actual PWL = 70 (distribution of the results from 1000 simulated projects).

Tests/Lot	Lots/Project	Tests/Project	Avg. Diff.	2.5%	97.5%
3	1	3	–1.0	–30.0	+44.0
	3	9	+0.3	–30.0	+29.6
	5	15	–0.5	–20.5	+21.7
	10	30	+0.2	–14.9	+14.5
	20	60	–0.0	–10.3	+10.8
	30	90	+0.1	–8.1	+8.7
	40	120	–0.1	–7.6	+6.9
	50	150	–0.1	–6.7	+6.4
	60	180	–0.1	–5.8	+6.0
	70	210	+0.0	–5.5	+5.6
	80	240	+0.2	–5.0	+5.5
	90	270	–0.0	–4.9	+5.2
5	1	5	–0.3	–30.0	+34.5
	3	15	–0.1	–19.2	+21.7
	5	25	+0.2	–15.1	+16.9
	10	50	+0.1	–10.3	+10.9
	20	100	–0.1	–7.7	+7.6
	30	150	+0.1	–6.2	+6.0
	40	200	–0.1	–5.5	+5.4
	50	250	+0.2	–4.8	+5.4
	60	300	+0.1	–4.5	+4.7
10	1	10	+0.5	–22.6	+27.3
	3	30	+0.2	–13.5	+13.7
	5	50	–0.0	–10.4	+10.5
	10	100	–0.0	–8.0	+7.5
	20	200	–0.0	–5.3	+5.1
	30	300	+0.0	–4.4	+4.5
	40	400	–0.0	–3.8	+3.5
	50	500	+0.0	–3.2	+3.4

Avg. Diff. = average difference between the average estimated PWL and the actual PWL for each simulated project

2.5% = value for which 2.5 percent of the simulated average differences were less than or equal

97.5% = value for which 97.5 percent of the simulated average differences were less than or equal

Table 11. Results of simulation analyses with actual PWL = 50 (distribution of the results from 1000 simulated projects).

Tests/Lot	Lots/Project	Tests/Project	Avg. Diff.	2.5%	97.5%
3	1	3	+0.0	–50.0	+50.0
	3	9	+0.1	–31.0	+29.6
	5	15	–0.5	–24.8	+22.6
	10	30	–0.2	–16.3	+16.2
	20	60	+0.2	–12.1	+11.7
	30	90	+0.1	–9.1	+9.4
	40	120	+0.0	–7.7	+8.3
	50	150	–0.0	–6.9	+7.1
	60	180	–0.0	–6.9	+6.7
	70	210	–0.0	–6.0	+6.0
	80	240	+0.1	–5.7	+6.2
	90	270	–0.0	–5.4	+5.6
5	1	5	–0.3	–41.0	+40.8
	3	15	+0.5	–21.4	+20.2
	5	25	+0.2	–16.6	+17.7
	10	50	–0.2	–12.2	+12.5
	20	100	–0.0	–8.5	+8.8
	30	150	–0.0	–6.5	+6.6
	40	200	+0.0	–5.7	+6.0
	50	250	–0.1	–5.0	+5.0
	60	300	+0.2	–4.9	+5.0
10	1	10	–0.0	–28.1	+29.0
	3	30	+0.2	–15.0	+14.7
	5	50	+0.0	–11.5	+10.5
	10	100	+0.2	–8.1	+8.2
	20	200	+0.0	–5.7	+5.5
	30	300	–0.1	–4.8	+4.3
	40	400	–0.0	–4.1	+3.9
	50	500	–0.1	–3.8	+3.6

Avg. Diff. = average difference between the average estimated PWL and the actual PWL for each simulated project

2.5% = value for which 2.5 percent of the simulated average differences were less than or equal

97.5% = value for which 97.5 percent of the simulated average differences were less than or equal

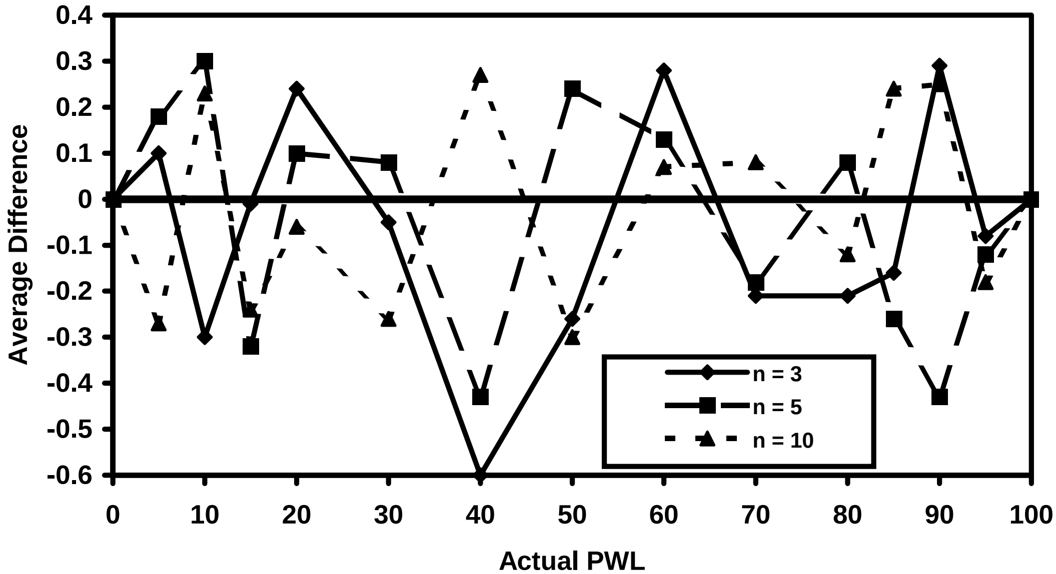


Figure 13. Plot of average difference of simulated PWL values versus actual PWL values for sample sizes = 3, 5, and 10.

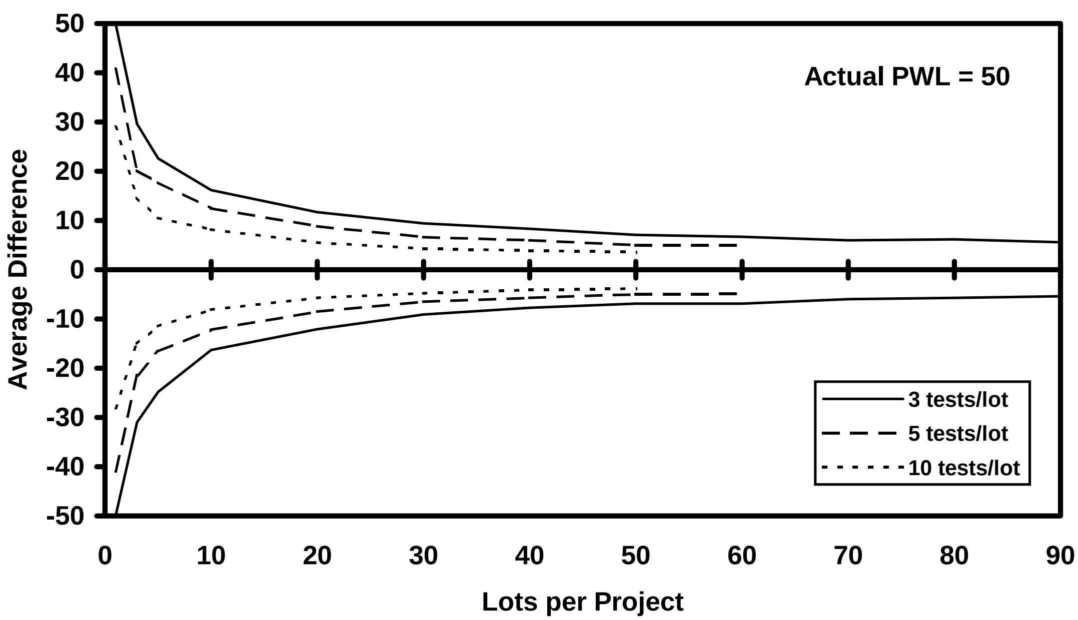


Figure 14a. Plots of the 95th percentile for the average estimated PWL minus the actual PWL at 50 versus the number of lots per project.

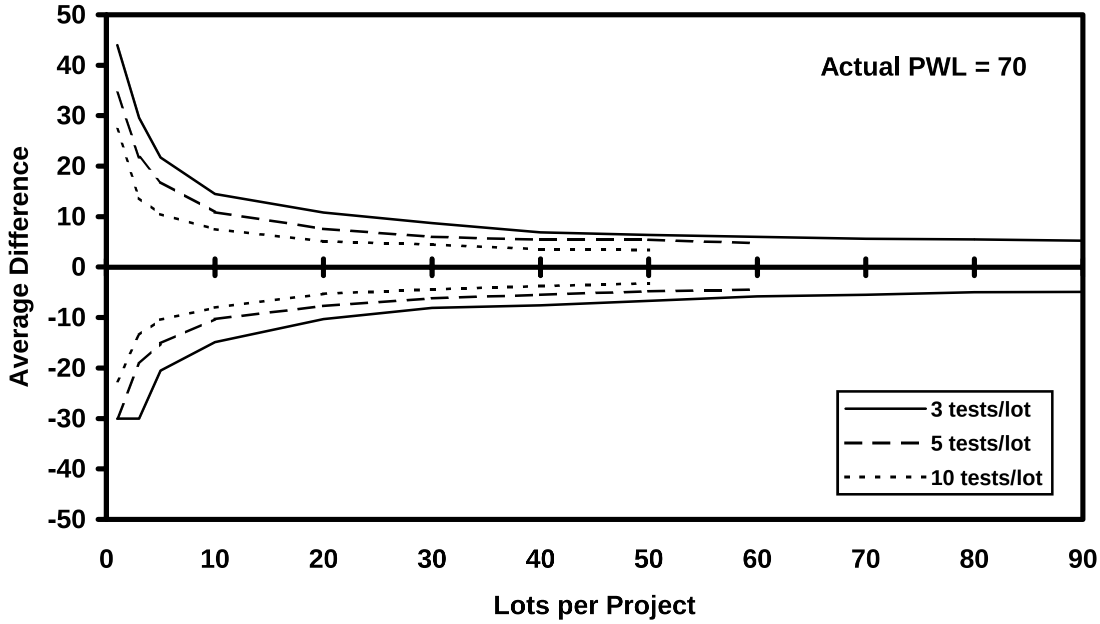


Figure 14b. Plots of the 95th percentile for the average estimated PWL minus the actual PWL at 70 versus the number of lots per project.

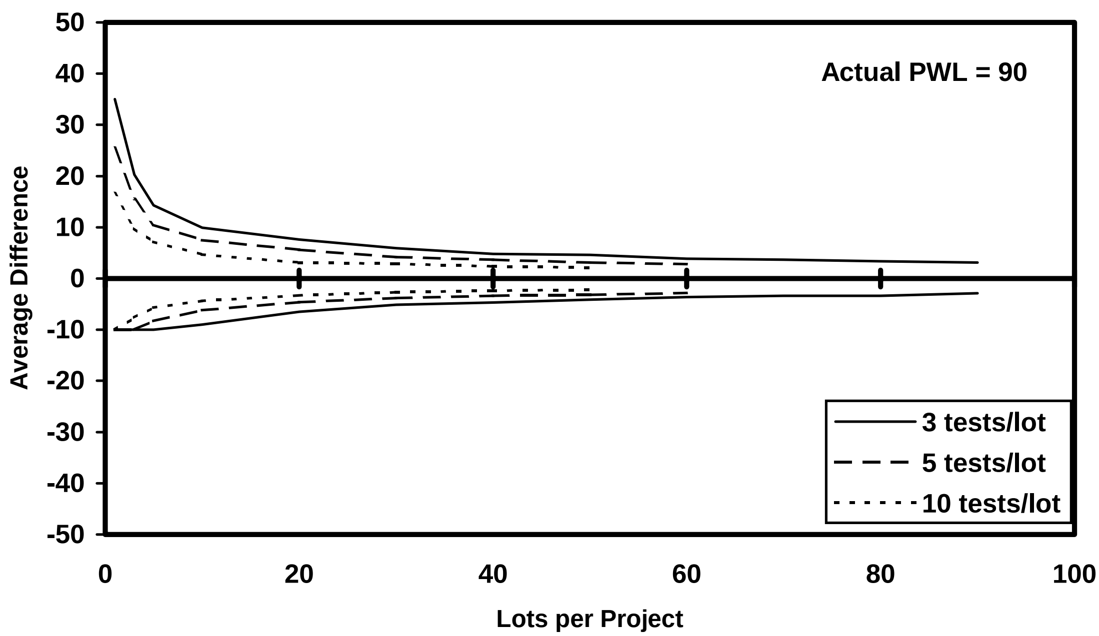


Figure 14c. Plots of the 95th percentile for the average estimated PWL minus the actual PWL at 90 versus the number of lots per project.

AAD EVALUATION

For specifications that have a target value, the average deviation from the target value has, in the past, sometimes been used as a means for determining the acceptability of the product. This approach can have the effect of encouraging the contractor to manipulate its process during the production of a lot. For example, if two test results in the morning are below the target value, there is a strong incentive for the contractor to increase the process mean in the afternoon in an effort to get two higher test results so that the average of the four tests for the lot will be near the target value. In essence, this acceptance approach encourages the contractor to increase process variability by making frequent adjustments to the process mean. To avoid the problem of overadjusting the process in response to early test results, the average *absolute* deviation from the target can be used for the acceptance decision.

The equation for calculating AAD is as follows:

$$AAD = \frac{\sum |X_i - T|}{n} \tag{4}$$

where:	X_i	=	individual test results
	T	=	target value
	n	=	number of tests per lot

Simulation Analyses

A computer simulation program was developed to evaluate the accuracy and precision of estimated AAD values. The program generates 1000 simulated lots. The mean, or average, estimated AAD value and the standard deviation for the estimated AAD values are then calculated.

The program can represent populations that are centered on the target value, but can also generate data for populations with means that are offset from the desired target. The actual AAD value for a population centered on the target is 0.798 multiplied by the standard deviation of the population. The actual AAD value increases as the population mean is offset from the target value. Without sacrificing generality, the simulation studies were performed for populations with standard deviation values of 1.00. For a population with a standard deviation other than 1.00, the simulated average AAD values would simply be multiplied by the standard deviation of the population in question.

For the study of AAD accuracy and precision, populations were simulated with offsets from the target that ranged from 0.00 to 2.50 in increments of 0.25. The populations from which the samples were generated were normal distributions with a standard deviation of 1.00.

The shape of the distribution of the sample AAD values varies with the sample size. For small mean offsets, the distribution is skewed to the right, with the degree of skewness determined by the sample size. For a sample size = 1, the peak of the distribution occurs at AAD values near zero. As the sample size increases, the value for the peak of the distribution increases and the degree of skewness decreases, because as the sample size increases, it becomes more and more difficult to have average AAD values near zero. This relationship is shown in table 12.

For a given sample size, the shape of the distribution of the sample AAD values also varies with the offset of the population mean from the target value. The skewness of the distribution stems from the fact that the deviations from the target cannot be below zero, thereby yielding distributions skewed away from zero. As the offset departs farther from the target, the distribution of the sample AAD values approaches symmetry since the deviation values are always greater than zero. This relationship is shown in table 13 for a sample size = 3 and for mean offsets = 0.50 to 2.50.

Table 12. Distributions of sample AAD values for a population centered on the target and for sample sizes = 1, 3, 5, and 10.

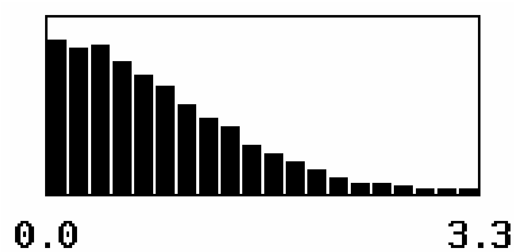
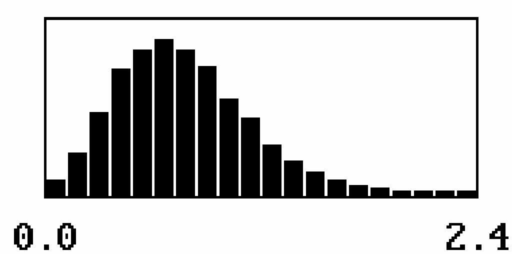
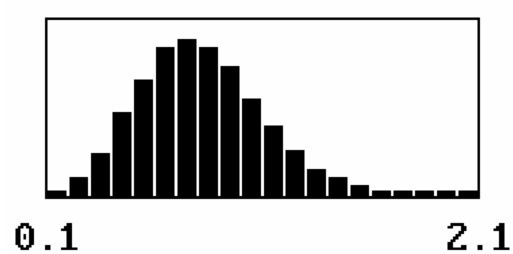
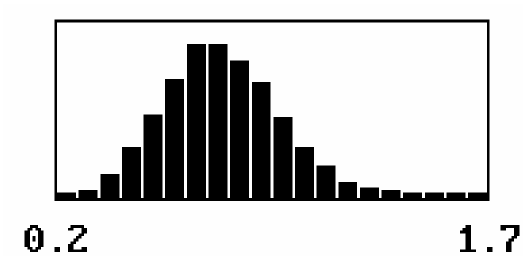
Sample Size, *n*	Distribution of Sample AAD Values
1	
3	
5	
10	

Table 13. Distributions of sample AAD values for sample size = 3 and population means offset from the target by 0.50, 1.00, 1.50, 2.00, and 2.50 standard deviations.

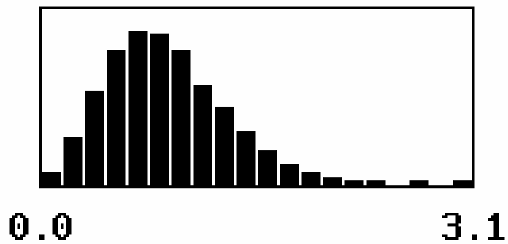
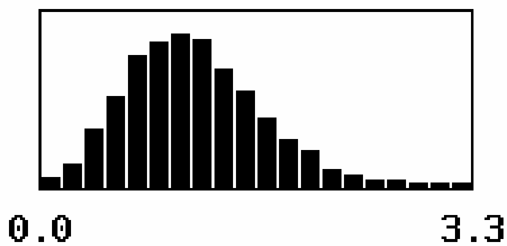
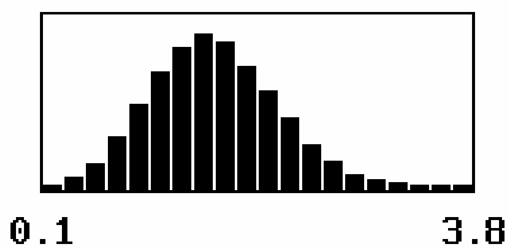
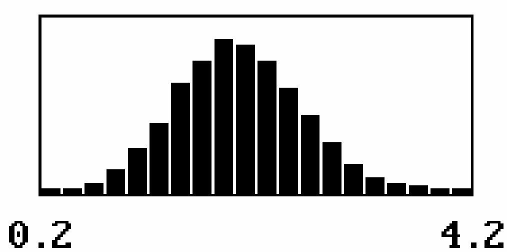
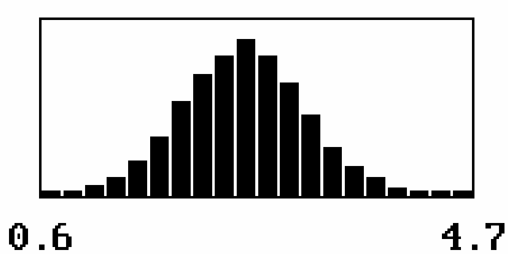
Mean Offset	Distribution of Sample AAD Values
0.50	
1.00	
1.50	
2.00	
2.50	

Table 14 presents a summary of the simulation results for AAD. The table includes sample sizes = 3, 5, and 10, and populations with means offset from the target by 0.00 to 2.50 standard deviations in increments of 0.25. The table also includes the actual AAD values for each population, which are based on the mean offset values. Finally, the table includes the average differences and the standard deviations for the estimated AAD values.

The results in table 14 show that the sample AAD is an unbiased estimator of the population AAD. It also shows that the standard deviation values decrease as the sample size increases and the standard deviation values increase as the offset from the target value increases. This is caused by the fact that as the offset increases, a deviation value of zero no longer presents a barrier, thereby allowing the deviations to spread evenly on both sides of the AAD value. These two trends are illustrated in figures 15 and 16.

CI EVALUATION

Conceptually, the CI is similar to AAD. While AAD uses the average of the absolute values of the individual deviations from the target value, the CI uses the squares of the individual deviations from the target value. The CI is also similar in concept to the standard deviation. The standard deviation is the root mean square of differences from the *mean*, whereas the CI is the root mean square of differences from a *target*. Like AAD, the CI discourages mid-lot process adjustments by not allowing positive and negative deviations from the target to cancel out one another. The CI is calculated as follows:

$$CI = \sqrt{\frac{\sum (X_i - T)^2}{n}} \qquad (5)$$

where: X_i = individual test results
T = target value
n = number of tests per lot

Simulation Analyses

A computer simulation program was developed to evaluate the accuracy and precision of estimated CI values. The program generated 10,000 simulated lots. The mean estimated CI value and the standard deviation for the estimated CI values are then calculated.

The simulation studies present the CI in terms of population standard deviation. Without sacrificing generality, the simulation studies were performed for populations with standard deviations of 1.00. For a population with a standard deviation other than 1.00, the simulated average CI values would simply be multiplied by the standard deviation of the population in question. The program can represent populations that have various CI values. If the population mean is centered on the target, then the CI is the same as the standard deviation. Therefore, the CI cannot be less than zero.

Table 14. Accuracy and precision for AAD estimates (based on the results of 10,000 simulated lots).

Tests per Lot, *n*	Offset of Population Mean From Target	Actual AAD of Population	Mean of Estimated AAD Values Minus Actual AAD	Standard Deviation of Estimated AAD Values
3	0.00	0.798	–0.003	0.3495
	0.25	0.823	+0.002	0.3624
	0.50	0.896	+0.003	0.3863
	0.75	1.012	–0.001	0.4223
	1.00	1.167	+0.005	0.4649
	1.25	1.351	–0.005	0.4911
	1.50	1.559	–0.001	0.5200
	1.75	1.782	+0.008	0.5496
	2.00	2.017	+0.004	0.5529
	2.25	2.258	+0.008	0.5660
	2.50	2.504	+0.003	0.5777
5	0.00	0.798	+0.002	0.2717
	0.25	0.823	+0.000	0.2758
	0.50	0.896	+0.002	0.2964
	0.75	1.012	–0.002	0.3257
	1.00	1.167	–0.005	0.3575
	1.25	1.351	+0.004	0.3822
	1.50	1.559	+0.000	0.4049
	1.75	1.782	+0.006	0.4216
	2.00	2.017	+0.002	0.4220
	2.25	2.258	+0.002	0.4386
	2.50	2.504	–0.004	0.4425
10	0.00	0.798	–0.002	0.1905
	0.25	0.823	+0.000	0.1979
	0.50	0.896	–0.003	0.2101
	0.75	1.012	+0.003	0.2293
	1.00	1.167	+0.003	0.2561
	1.25	1.351	–0.001	0.2693
	1.50	1.559	–0.002	0.2848
	1.75	1.782	–0.002	0.2978
	2.00	2.017	+0.001	0.3033
	2.25	2.258	–0.003	0.3131
	2.50	2.504	–0.004	0.3152

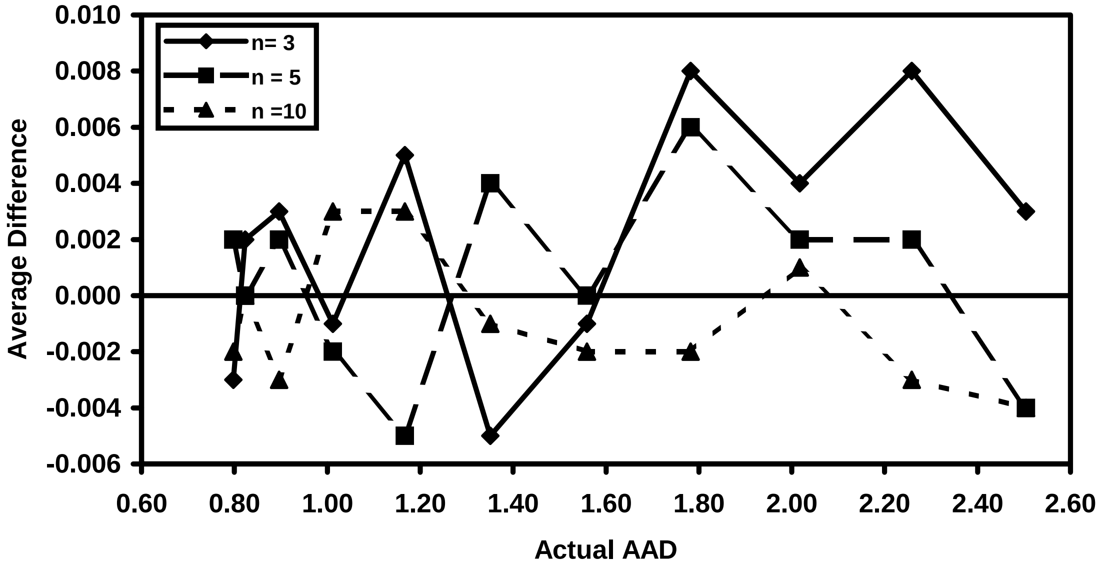


Figure 15. Plot of average difference of simulated AAD values versus actual AAD values for sample sizes = 3, 5, and 10.

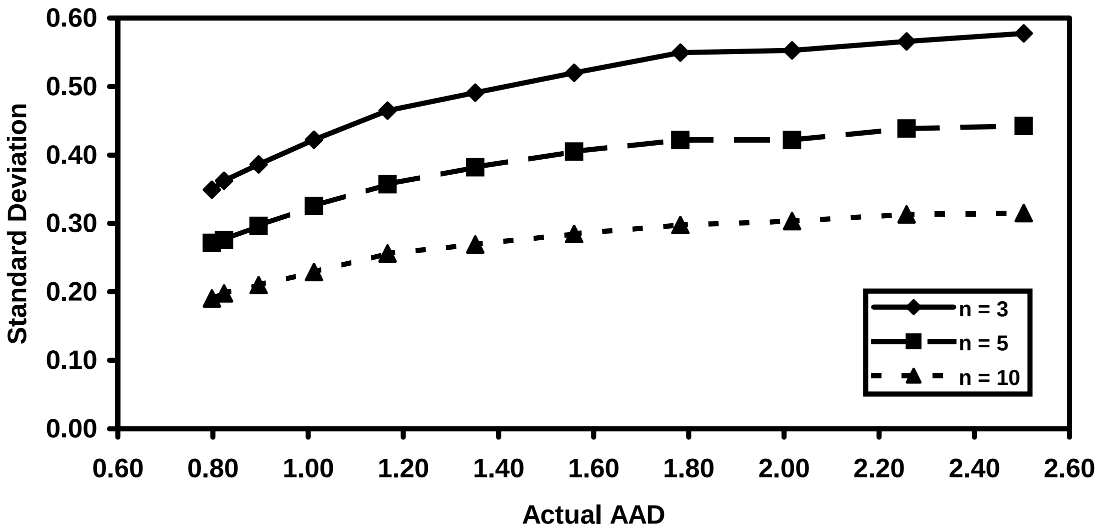


Figure 16. Plot of how the standard deviation of AAD estimates varies with the population AAD value.

For the study of CI accuracy and precision, populations were simulated with CI values ranging from 1.00 to 3.00. The populations from which the samples were generated were normal distributions with a standard deviation of 1.00.

The shape of the distribution of the sample CI values varies with the sample size in a similar manner to that of AAD. For small CI values, the distribution is skewed to the right, with the degree of skewness determined by the sample size. As the sample size increases, the degree of skewness decreases, because as the sample size increases, it becomes more and more difficult to have average CI values near zero.

Similar to that of AAD, for a given sample size, the shape of the distribution of the sample CI values varies with the CI of the population. For small values of CI, the skewness of the distribution stems from the fact that the squared deviations from the target cannot be below zero, thereby yielding distributions skewed away from zero. As the CI increases, meaning that the population mean is farther from the target, the distribution approaches symmetry since the squared deviation values are always greater than zero.

The results in table 15 show that the sample CI appears to be a slightly biased estimator of the population CI since the average of the CI estimates is always lower than the population CI. This bias decreases as the sample size increases. The table also shows that the standard deviation values decrease as the sample size increases and the standard deviation values increase as the CI value for the population increases. These results are illustrated in figures 17 and 18.

CONCLUSION

Both the PWL and AAD estimators were unbiased and exhibited similar trends with respect to the variability of their individual estimates. The CI performed essentially the same as AAD, except that it appeared to be a slightly biased estimator. Since the CI offered no benefits compared to AAD and since it appeared slightly biased, it was decided to eliminate the CI and to concentrate further study on only PWL and AAD.

Table 15. Accuracy and precision for CI estimates (based on the results of 10,000 simulated lots).

Tests per Lot, *n*	Actual CI of Population	Mean of Estimated CI Values Minus Actual CI	Standard Deviation of Estimated CI Values
3	1.00	–0.09	0.39
	1.10	–0.07	0.43
	1.25	–0.10	0.44
	1.40	–0.07	0.41
	1.50	–0.08	0.49
	1.75	–0.08	0.52
	2.00	–0.07	0.54
	2.25	–0.05	0.53
	2.50	–0.05	0.56
	3.00	–0.05	0.56
5	1.00	–0.05	0.31
	1.10	–0.05	0.34
	1.25	–0.04	0.37
	1.40	–0.06	0.38
	1.50	–0.08	0.39
	1.75	–0.02	0.39
	2.00	–0.06	0.43
	2.25	–0.01	0.43
	2.50	–0.04	0.44
	3.00	–0.03	0.44
10	1.00	–0.01	0.22
	1.10	–0.03	0.24
	1.25	–0.03	0.27
	1.40	–0.03	0.27
	1.50	–0.03	0.27
	1.75	–0.02	0.29
	2.00	–0.02	0.30
	2.25	–0.01	0.31
	2.50	0.00	0.30
	3.00	–0.02	0.30

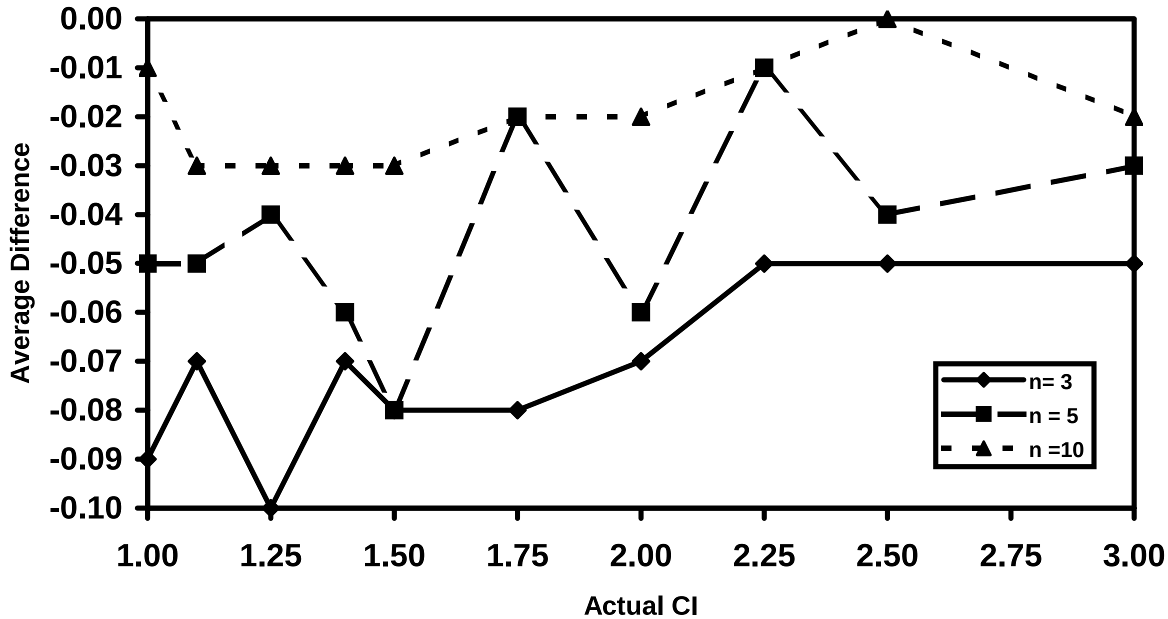


Figure 17. Plot of average difference of simulated CI values versus actual CI values for sample sizes = 3, 5, and 10.

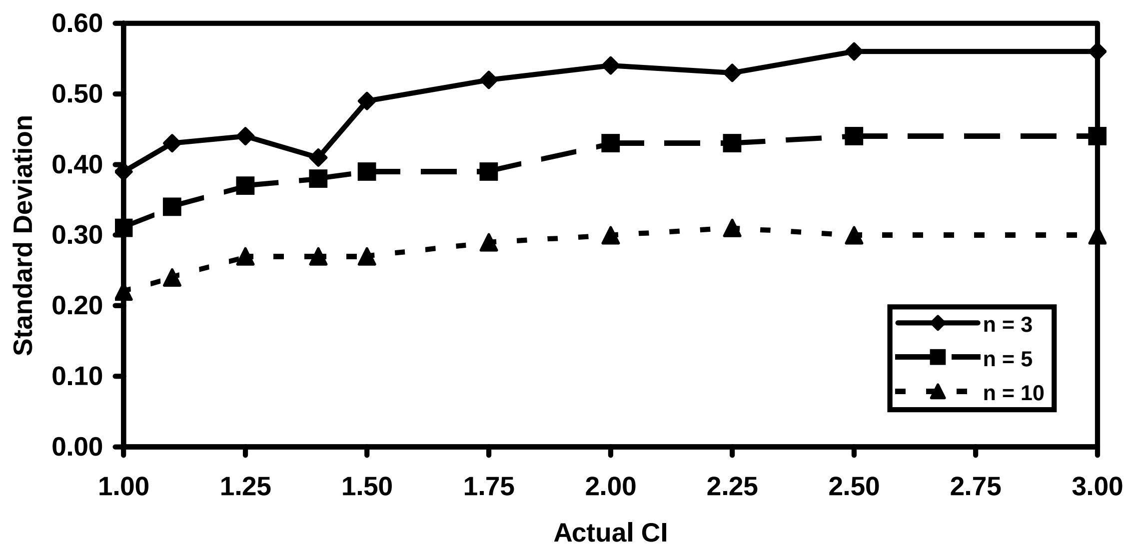


Figure 18. Plot of how the standard deviation of CI estimates varies with the population CI value.

6. QUALITY MEASURES: NORMALITY ASSUMPTION

INTRODUCTION

As noted in chapter 5, the PWL and AAD quality measures were selected for additional analyses, while the CI was eliminated from further study. The use of each of these measures is based on the assumption that the population that is being sampled is normally distributed. To determine how robust these measures are, computer simulation was used to evaluate how they performed when the sampled population was not normally distributed.

The most likely departures from normality would be skewed or bimodal distribution of data. Computer simulation programs were developed to evaluate the performance of the PWL and AAD estimators under each of these situations. The following sections of this chapter present the results for analyses conducted using these simulation programs.

SKEWED DISTRIBUTIONS

Skewed distributions usually occur because of some physical boundary that comes into play for a particular characteristic. For example, the percent passing a sieve for gradation analysis cannot exceed 100 percent. Thus, if the average percent passing is near 100, say 95 percent, it is possible to have greater spread on the low side of the average than on the high side. Another barrier might be pavement thickness, which cannot be less than zero. These cases are similar in concept to the discussions in chapter 5 regarding why the distributions of sample AAD and CI values have skewed distributions. The following sections present the results of computer simulation analyses on the effects of skewness on the estimates for PWL and AAD.

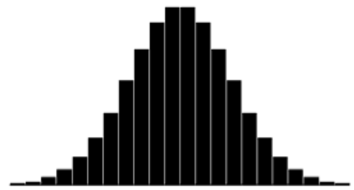
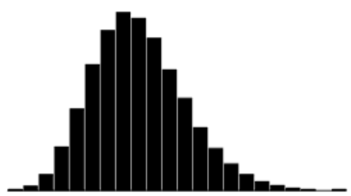
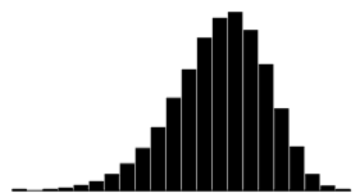
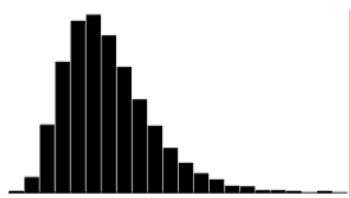
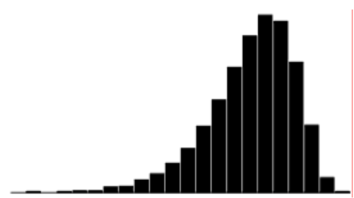
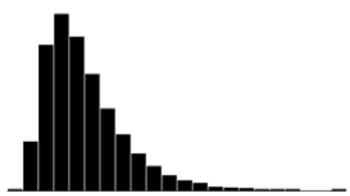
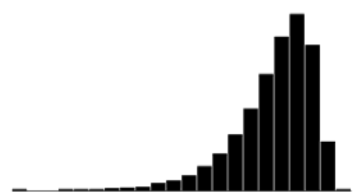
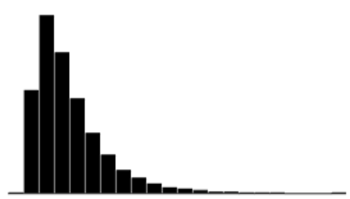
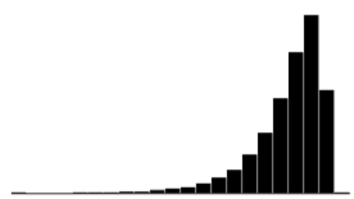
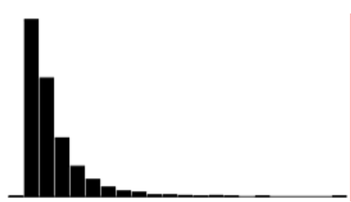
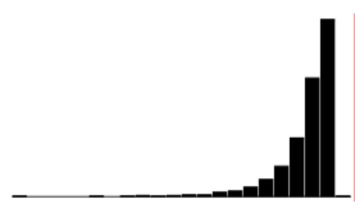
PWL Evaluation

Two computer simulation programs were developed to investigate the performance of the PWL estimating procedures when working with a skewed distribution. One program is for use with one-sided specifications, while the other is for two-sided specifications. Skewness is measured by a skewness coefficient that equals zero for a symmetrical distribution and which increases as the distribution becomes more highly skewed. The skewness for a data set can be calculated from the following equation:

$$skewness = \frac{n}{(n-1)(n-2)} \sum \left(\frac{x_i - \overline{X}}{s} \right)^3 \qquad (6)$$

A skewness value of zero indicates that the distribution of data is symmetrical. The skewness can be either positive or negative. Positive skewness indicates a distribution with the long tail to the right, while negative skewness indicates that the long tail is to the left. Table 16 illustrates distributions with various levels of skewness.

Table 16. Distributions with various levels of skewness.

Skewness Coefficient	Positive	Negative
0.0		
0.5		
1.0		
1.5		
2.0		
3.0		

Both of the programs begin with a data set that is normally distributed. An exponent, entered by the user, is then used to transform these data to the skewed distribution desired.

One-Sided Limit: Computer simulation was used to evaluate the PWL estimates for populations with skewness coefficients of 0, +0.5, –0.5, +1.0, –1.0, +1.5, –1.5, +2.0, –2.0, +2.5, –2.5, +3.0, and –3.0. Sample sizes = 3, 5, and 10 were used to estimate the PWL values for populations with each of these skewness coefficients. In each analysis, 10,000 samples of the appropriate size were generated from a population with one of the skewness coefficients. The average bias and variability were then determined for the 10,000 PWL estimates. The results of these simulations are shown in tables 17 through 19 for sample sizes = 3, 5, and 10, respectively.

For efficiency, the simulation program calculates the bias and standard deviation only for positive skewness coefficients, it then reverses these numbers and their signs to obtain the values for negative skewness (i.e., the bias for 90 PWL with positive skewness is equal to –1 times the bias for 10 PWL with negative skewness). This is apparent in tables 17 through 19.

Figures 19 through 21 graphically present the bias results from tables 17 through 19, respectively. These figures plot the bias versus the actual PWL value for each of the levels of skewness. Figures 19 through 21 present the results for samples sizes = 3, 5, and 10, respectively.

The trends are obvious. The greater the skewness, the greater the bias in the PWL estimate. This is logical since the distributions deviate more from normality (on which the PWL calculation is based) as the skewness increases. The values for positive and negative skewness are mirror images of each other.

Table 20 presents a portion of the bias results from tables 17 through 19. The results are sorted by the amount of skewness, with skewness coefficients of 1.0, 2.0, and 3.0 included. Figure 22 presents the bias results from table 20 plotted by sample sizes = 3, 5, and 10 for equal levels of skewness. It should be noted that for negative skewness, the magnitude of the biases would be the same; however, the plots would be flipped vertically about the bias value of zero.

Interestingly, figure 22 shows that the larger the sample size, the greater the bias of the estimate. This probably stems from the fact that the larger sample does a better job of estimating PWL for the assumed symmetrical normal distribution. This leads to a greater bias since the population is actually a skewed distribution rather than a symmetrical normal distribution.

Table 17. Bias in estimating PWL for various skewness coefficients and one-sided limits (3 tests per lot and 10,000 simulated lots).

PWL	+0.0	+0.5	+1.0	+1.5	+2.0	+2.5	+3.0
99	+0.04	–0.35	–1.26	–2.33	–3.57	–4.96	–6.57
95	–0.04	–0.60	–1.27	–2.35	–3.14	–4.31	–5.43
90	–0.01	–0.33	–1.04	–1.63	–2.11	–2.88	–3.65
80	+0.35	+0.23	+0.25	–0.23	–0.41	–0.45	–0.53
70	+0.64	+0.72	+0.77	+1.42	+0.81	+1.41	+1.39
60	–0.12	+0.11	+1.21	+1.69	+2.04	+2.53	+2.82
50	–0.35	+0.61	+1.61	+2.13	+2.95	+3.20	+3.67
40	+0.14	+0.83	+0.81	+2.03	+2.94	+3.13	+3.76
30	–0.10	+0.61	+1.11	+1.63	+2.07	+2.22	+2.95
20	–0.31	+0.07	+0.54	+0.90	+1.77	+1.43	+1.87
10	+0.11	–0.19	–0.10	–0.16	+0.33	+0.51	+0.23
5	+0.02	–0.21	–0.05	–0.27	–0.33	–0.17	–0.10
1	+0.06	–0.12	–0.19	–0.22	–0.21	–0.20	–0.22
PWL	**–0.0**	**–0.5**	**–1.0**	**–1.5**	**–2.0**	**–2.5**	**–3.0**
99	–0.06	+0.12	+0.19	+0.22	+0.21	+0.20	+0.22
95	–0.02	+0.21	+0.05	+0.27	+0.33	+0.17	+0.10
90	–0.11	+0.19	+0.10	+0.16	–0.33	–0.51	–0.23
80	+0.31	–0.07	–0.54	–0.90	–1.77	–1.43	–1.87
70	+0.10	–0.61	–1.11	–1.63	–2.07	–2.22	–2.95
60	–0.14	–0.83	–0.81	–2.03	–2.94	–3.13	–3.76
50	+0.35	–0.61	–1.61	–2.13	–2.95	–3.20	–3.67
40	+0.12	–0.11	–1.21	–1.69	–2.04	–2.53	–2.82
30	–0.64	–0.72	–0.77	–1.42	–0.81	–1.41	–1.39
20	–0.35	–0.23	–0.25	+0.23	+0.41	+0.45	+0.53
10	+0.01	+0.33	+1.04	+1.63	+2.11	+2.88	+3.65
5	+0.04	+0.60	+1.27	+2.35	+3.14	+4.31	+5.43
1	–0.04	+0.35	+1.26	+2.33	+3.57	+4.96	+6.57

Note: The values shown in the table are calculated by subtracting the actual population PWL value from the average of the 10,000 estimated lot PWL values.

Table 18. Bias in estimating PWL for various skewness coefficients and one-sided limits (5 tests per lot and 10,000 simulated lots).

PWL	+0.0	+0.5	+1.0	+1.5	+2.0	+2.5	+3.0
99	+0.03	–0.78	–2.27	–4.31	–6.27	–8.22	–9.96
95	–0.20	–0.97	–2.34	–4.05	–5.32	–6.89	–8.26
90	–0.03	–0.56	–1.65	–2.54	–3.72	–4.76	–5.68
80	–0.20	+0.22	+0.15	+0.14	+0.01	–0.40	–0.65
70	–0.46	+1.04	+1.79	+2.37	+2.74	+3.04	+3.14
60	–0.19	+1.24	+2.54	+3.58	+4.90	+5.36	+5.47
50	–0.02	+1.15	+3.12	+4.42	+5.54	+6.50	+7.17
40	–0.08	+1.44	+3.24	+4.68	+5.42	+6.36	+7.30
30	–0.03	+1.17	+2.13	+3.02	+4.21	+5.29	+6.13
20	+0.39	+0.16	+0.97	+1.58	+2.26	+2.82	+3.81
10	–0.13	–0.44	–0.47	–0.12	+0.35	+0.43	+0.80
5	–0.04	–0.77	–0.86	–0.70	–0.77	–0.65	–0.57
1	0.00	–0.30	–0.40	–0.47	–0.45	–0.46	–0.36
PWL	**–0.0**	**–0.5**	**–1.0**	**–1.5**	**–2.0**	**–2.5**	**–3.0**
99	0.00	+0.30	+0.40	+0.47	+0.45	+0.46	+0.36
95	+0.04	+0.77	+0.86	+0.70	+0.77	+0.65	+0.57
90	+0.13	+0.44	+0.47	+0.12	–0.35	–0.43	–0.80
80	–0.39	–0.16	–0.97	–1.58	–2.26	–2.82	–3.81
70	+0.03	–1.17	–2.13	–3.02	–4.21	–5.29	–6.13
60	+0.08	–1.44	–3.24	–4.68	–5.42	–6.36	–7.30
50	+0.02	–1.15	–3.12	–4.42	–5.54	–6.50	–7.17
40	+0.19	–1.24	–2.54	–3.58	–4.90	–5.36	–5.47
30	+0.46	–1.04	–1.79	–2.37	–2.74	–3.04	–3.14
20	+0.20	–0.22	–0.15	–0.14	–0.01	+0.40	+0.65
10	+0.03	+0.56	+1.65	+2.54	+3.72	+4.76	+5.68
5	+0.20	+0.97	+2.34	+4.05	+5.32	+6.89	+8.26
1	–0.03	+0.78	+2.27	+4.31	+6.27	+8.22	+9.96

Note: The values shown in the table are calculated by subtracting the actual population PWL value from the average of the 10,000 estimated lot PWL values.

Table 19. Bias in estimating PWL for various skewness coefficients and one-sided limits (10 tests per lot and 10,000 simulated lots).

PWL	+0.0	+0.5	+1.0	+1.5	+2.0	+2.5	+3.0
99	+0.01	–1.15	–3.21	–5.77	–8.28	–10.61	–12.63
95	+0.01	–1.35	–3.24	–5.24	–7.32	–9.02	–10.47
90	+0.08	–0.76	–2.14	–3.55	–5.08	–6.39	–7.64
80	0.00	+0.26	+0.41	–0.05	–0.54	–0.98	–1.69
70	–0.29	+1.66	+2.49	+3.15	+3.45	+3.46	+3.49
60	–0.18	+2.08	+4.11	+5.47	+6.41	+7.27	+7.72
50	+0.04	+2.42	+4.76	+6.47	+8.27	+9.27	+10.17
40	+0.02	+2.48	+4.54	+6.33	+8.29	+9.60	+11.11
30	–0.05	+1.43	+3.63	+5.34	+7.02	+8.08	+9.75
20	+0.09	+0.46	+1.60	+2.99	+3.79	+5.03	+6.29
10	+0.06	–0.53	–0.64	–0.22	+0.03	+0.20	+1.06
5	–0.04	–0.78	–1.18	–1.29	–1.12	–0.97	–0.72
1	+0.01	–0.44	–0.64	–0.64	–0.69	–0.63	–0.60
PWL	**–0.0**	**–0.5**	**–1.0**	**–1.5**	**–2.0**	**–2.5**	**–3.0**
99	–0.01	+0.44	+0.64	+0.64	+0.69	+0.63	+0.60
95	+0.04	+0.78	+1.18	+1.29	+1.12	+0.97	+0.72
90	–0.06	+0.53	+0.64	+0.22	–0.03	–0.20	–1.06
80	–0.09	–0.46	–1.60	–2.99	–3.79	–5.03	–6.29
70	+0.05	–1.43	–3.63	–5.34	–7.02	–8.08	–9.75
60	–0.02	–2.48	–4.54	–6.33	–8.29	–9.60	–11.11
50	–0.04	–2.42	–4.76	–6.47	–8.27	–9.27	–10.17
40	+0.18	–2.08	–4.11	–5.47	–6.41	–7.27	–7.72
30	+0.29	–1.66	–2.49	–3.15	–3.45	–3.46	–3.49
20	+0.00	–0.26	–0.41	+0.05	+0.54	+0.98	+1.69
10	–0.08	+0.76	+2.14	+3.55	+5.08	+6.39	+7.64
5	–0.01	+1.35	+3.24	+5.24	+7.32	+9.02	+10.47
1	–0.01	+1.15	+3.21	+5.77	+8.28	+10.61	+12.63

Note: The values shown in the table are calculated by subtracting the actual population PWL value from the average of the 10,000 estimated lot PWL values.

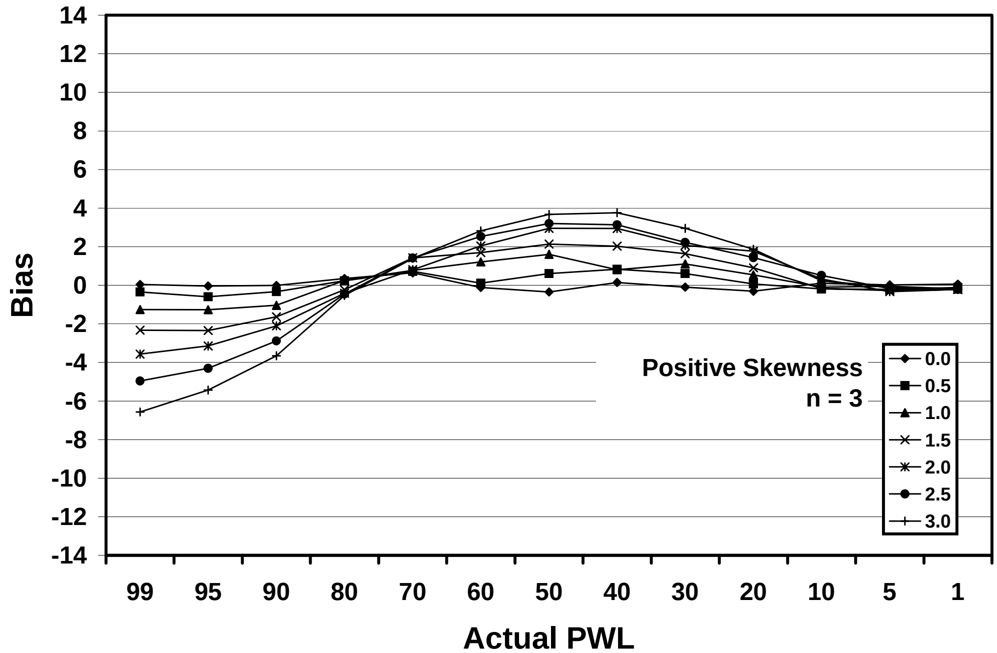


Figure 19a. Plots of bias versus actual PWL for 10,000 simulated lots with 3 tests per lot and one-sided limits showing positive skewness.

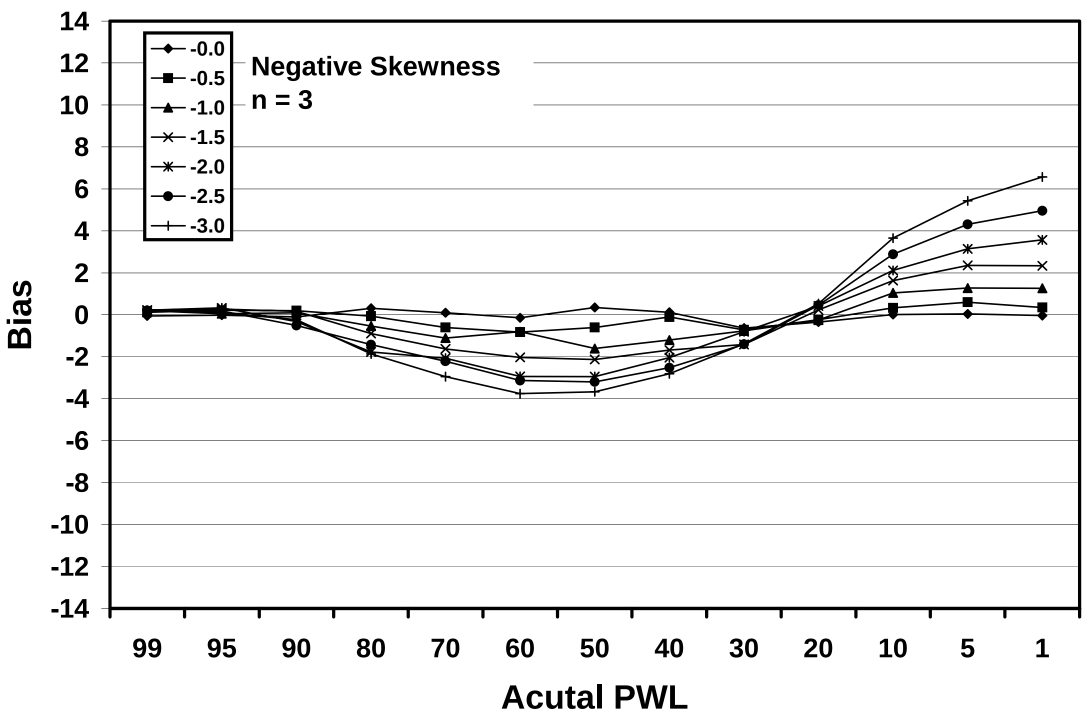


Figure 19b. Plots of bias versus actual PWL for 10,000 simulated lots with 3 tests per lot and one-sided limits showing negative skewness.

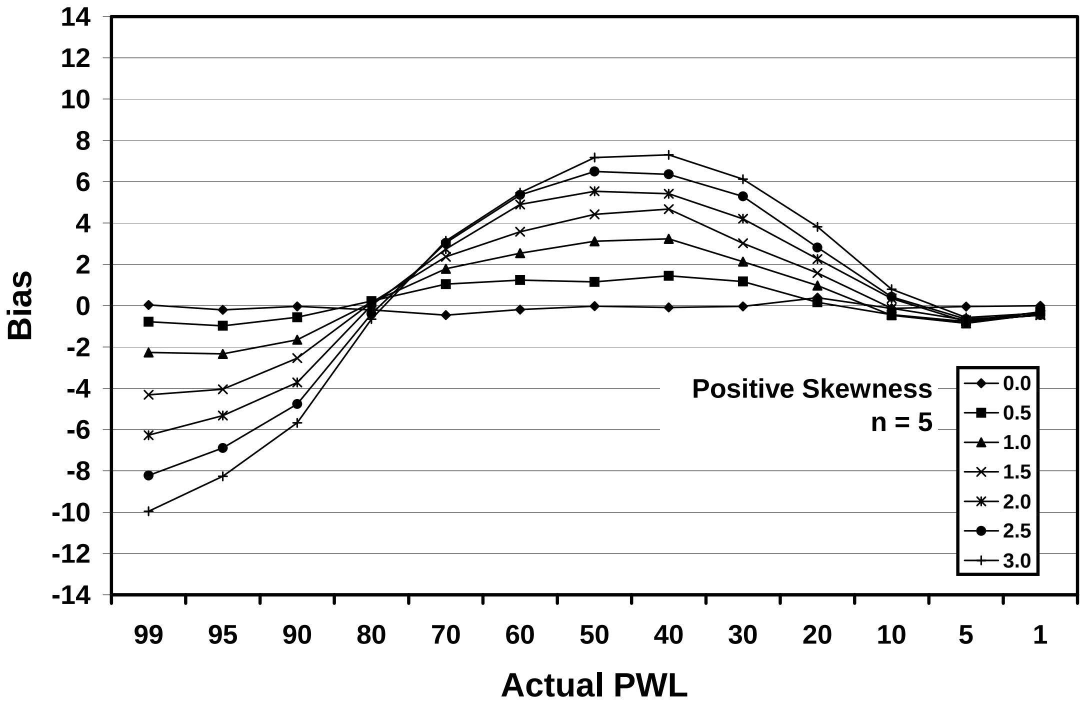


Figure 20a. Plots of bias versus actual PWL for 10,000 simulated lots with 5 tests per lot and one-sided limits showing positive skewness.

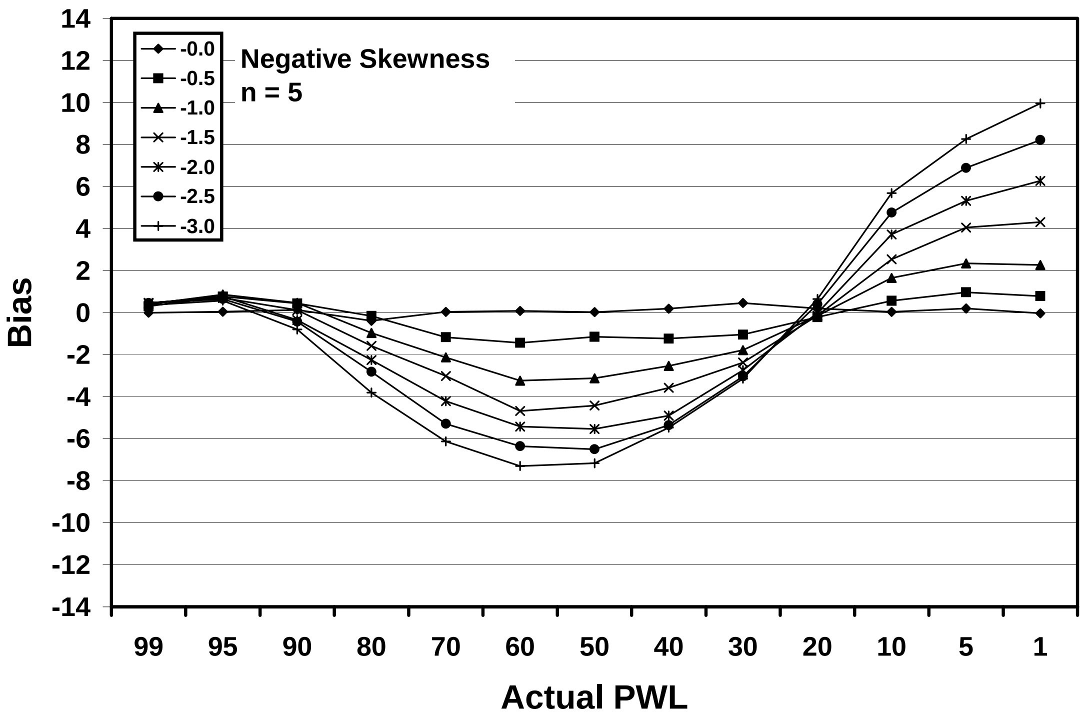


Figure 20b. Plots of bias versus actual PWL for 10,000 simulated lots with 5 tests per lot and one-sided limits showing negative skewness.

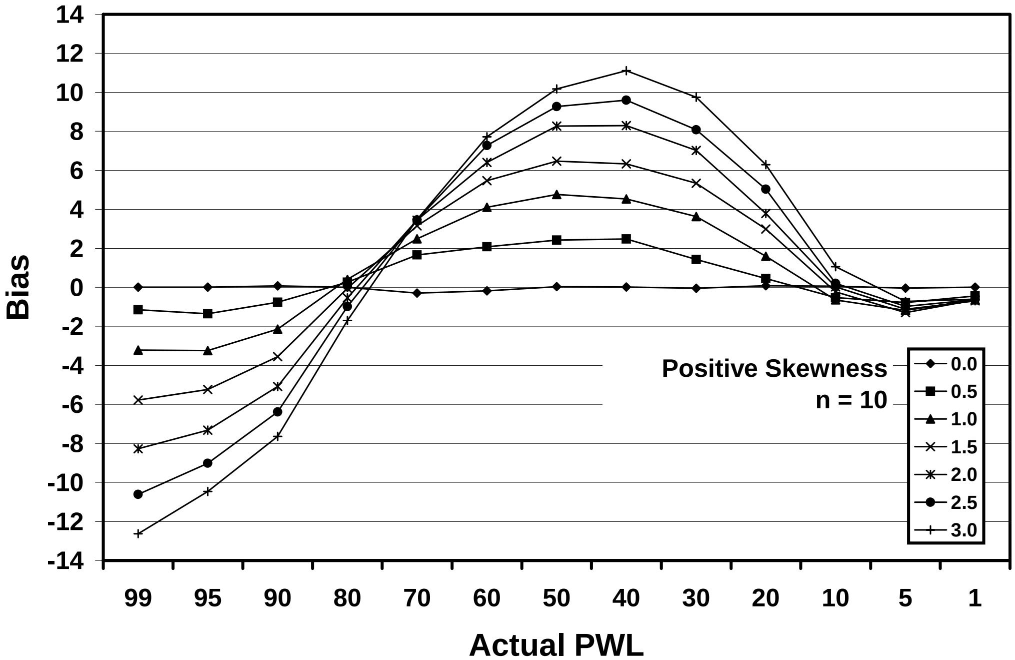


Figure 21a. Plots of bias versus actual PWL for 10,000 simulated lots with 10 tests per lot and one-sided limits with positive skewness.

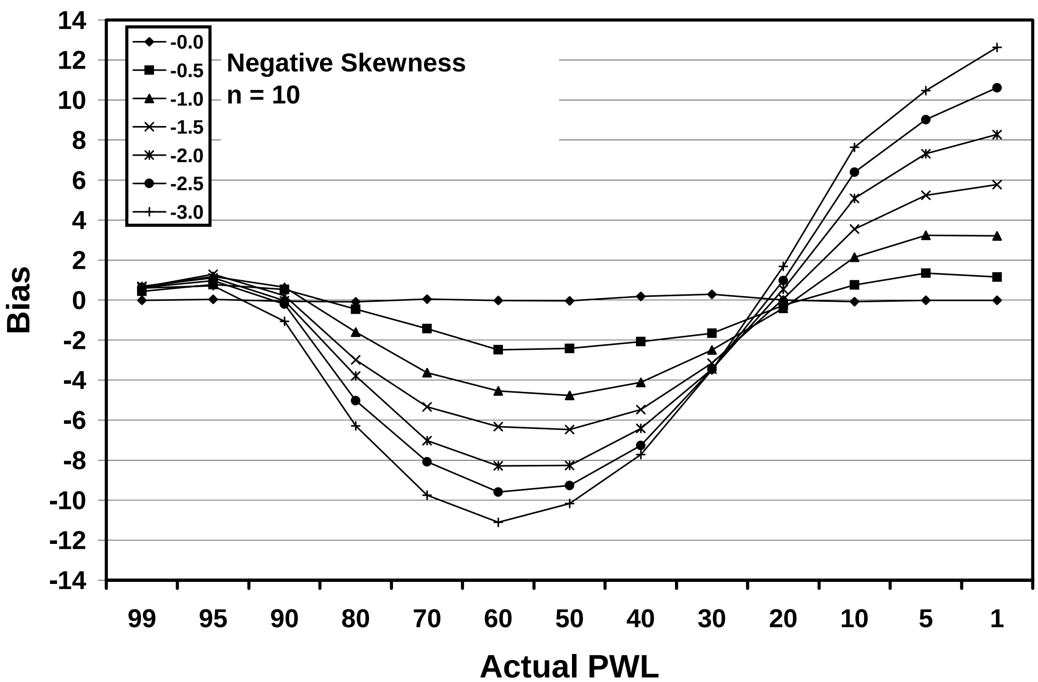


Figure 21b. Plots of bias versus actual PWL for 10,000 simulated lots with 10 tests per lot and one-sided limits with negative skewness.

Table 20. Bias in estimating PWL for various skewness coefficients, sample sizes, and one-sided limits (10,000 simulated lots).

PWL	Skewness = +1.0			Skewness = +2.0			Skewness = +3.0		
	***n* = 3**	***n* = 5**	***n* = 10**	***n* = 3**	***n* = 5**	***n* = 10**	***n* = 3**	***n* = 5**	***n* = 10**
99	–1.26	–2.27	–3.21	–3.57	–6.27	–8.28	–6.57	–9.96	–12.63
95	–1.27	–2.34	–3.24	–3.14	–5.32	–7.32	–5.43	–8.26	–10.47
90	–1.04	–1.65	–2.14	–2.11	–3.72	–5.08	–3.65	–5.68	–7.64
80	+0.25	+0.15	+0.41	–0.41	+0.01	–0.54	–0.53	–0.65	–1.69
70	+0.77	+1.79	+2.49	+0.81	+2.74	+3.45	+1.39	+3.14	+3.49
60	+1.21	+2.54	+4.11	+2.04	+4.90	+6.41	+2.82	+5.47	+7.72
50	+1.61	+3.12	+4.76	+2.95	+5.54	+8.27	+3.67	+7.17	+10.17
40	+0.81	+3.24	+4.54	+2.94	+5.42	+8.29	+3.76	+7.30	+11.11
30	+1.11	+2.13	+3.63	+2.07	+4.21	+7.02	+2.95	+6.13	+9.75
20	+0.54	+0.97	+1.60	+1.77	+2.26	+3.79	+1.87	+3.81	+6.29
10	–0.10	–0.47	–0.64	+0.33	+0.35	+0.03	+0.23	+0.80	+1.06
5	–0.05	–0.86	–1.18	–0.33	–0.77	–1.12	–0.10	–0.57	–0.72
1	–0.19	–0.40	–0.64	–0.21	–0.45	–0.69	–0.22	–0.36	–0.60

Note: The values shown in the table are calculated by subtracting the actual population PWL value from the average of the 10,000 estimated lot PWL values.

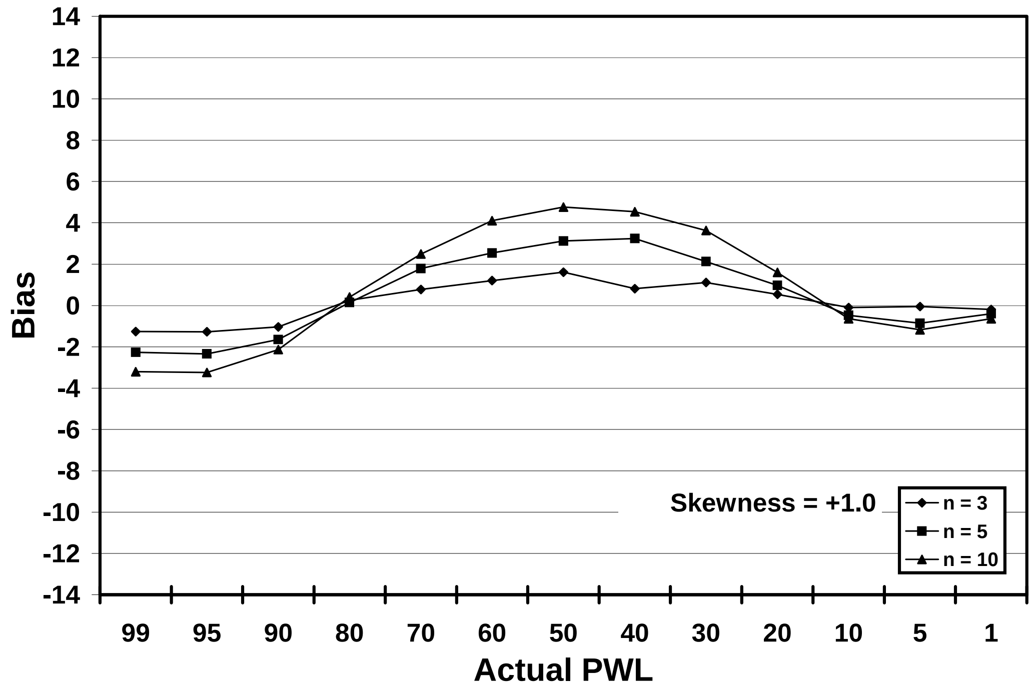


Figure 22a. Plot of bias versus actual PWL for 10,000 simulated lots with various tests per lot and one-sided limits with +1 skewness.

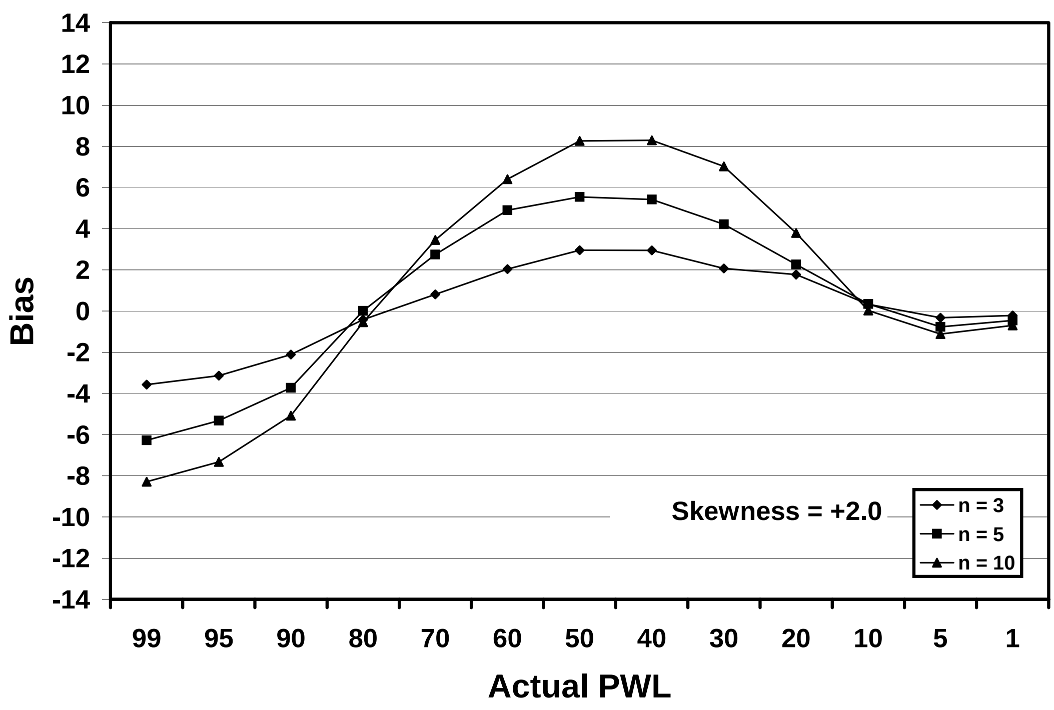


Figure 22b. Plot of bias versus actual PWL for 10,000 simulated lots with various tests per lot and one-sided limits with +2 skewness.

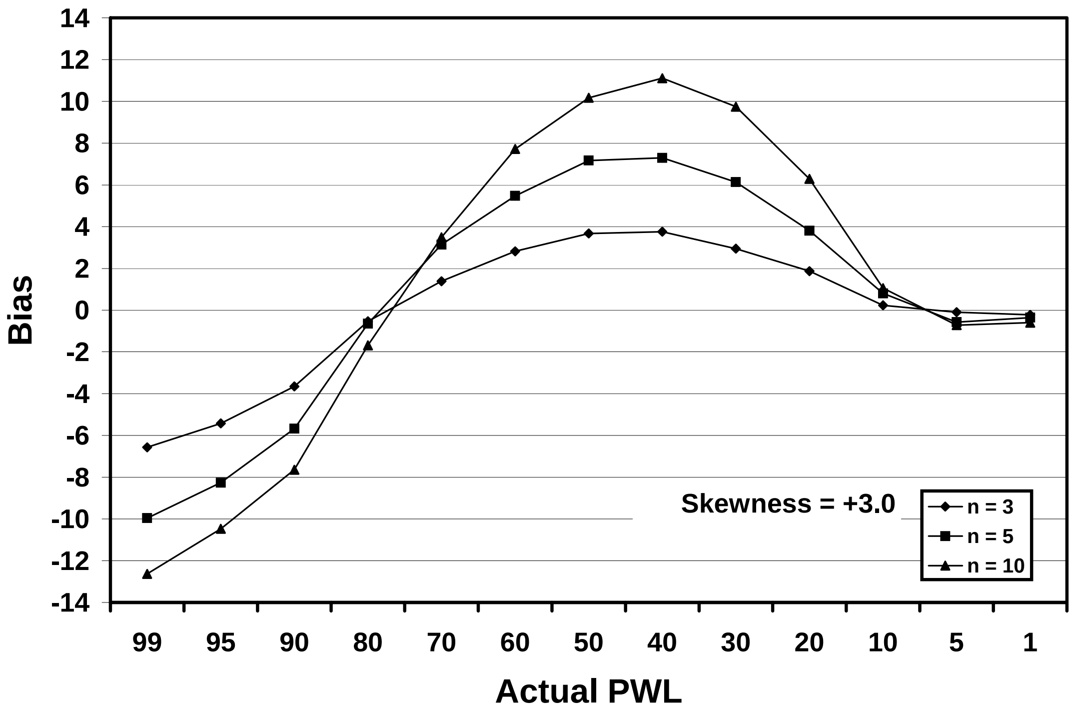


Figure 22c. Plot of bias versus actual PWL for 10,000 simulated lots with various tests per lot and one-sided limits with +3 skewness.

Two-Sided Limits: Computer simulation was used to study a number of different cases of skewed populations when there were two-sided specification limits. Skewness coefficients of 0, 0.5, 1.0, 1.5, 2.0, 2.5, and 3.0 were considered along with sample sizes of n = 3, 5, and 10. Because of the relative ease of use of PD for two-sided limits (in comparison to using PWL), the simulations were based on PD. Since it has been shown earlier in this report that PD and PWL are equivalent, but complementary, measures, the discussions of two-sided limits are based on PD rather than PWL.

With two-sided specification limits, there are many ways in which a given PD value can be divided between being outside the lower and upper limits. The program (SKEWBIAS2H) that was developed simulates a number of different divisions of the PD areas. The divisions can be noted as PD_L and PD_U, where PD_L is the percent defective below the lower specification limit and PD_U is the percent defective above the upper specification limit. The ratio of the division of PD_L and PD_U used are 8/0, 7/1, 6/2, 5/3, 4/4, 3/5, 2/6, 1/7, and 0/8. The diagrams presented in figure 23 illustrate the PD divisions used.

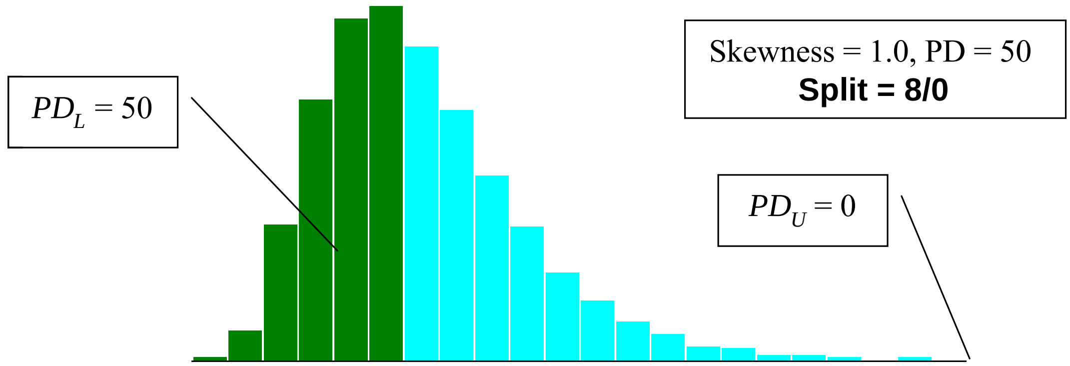


Figure 23a. Illustration of the divisions that SKEWBIAS2H uses to calculate bias in the PWL estimate for two-sided specification limits (skewness coefficient = +1.0, split=8/0).

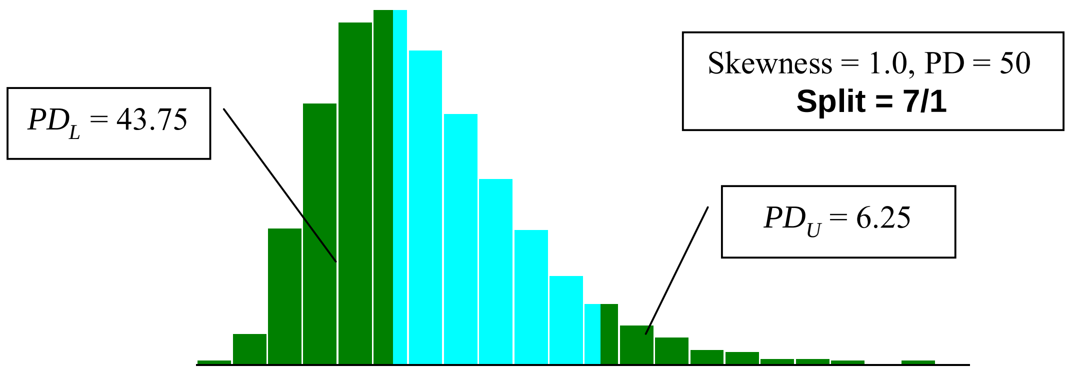


Figure 23b. Illustration of the divisions that SKEWBIAS2H uses to calculate bias in the PWL estimate for two-sided specification limits (skewness coefficient = +1.0, split=7/1).

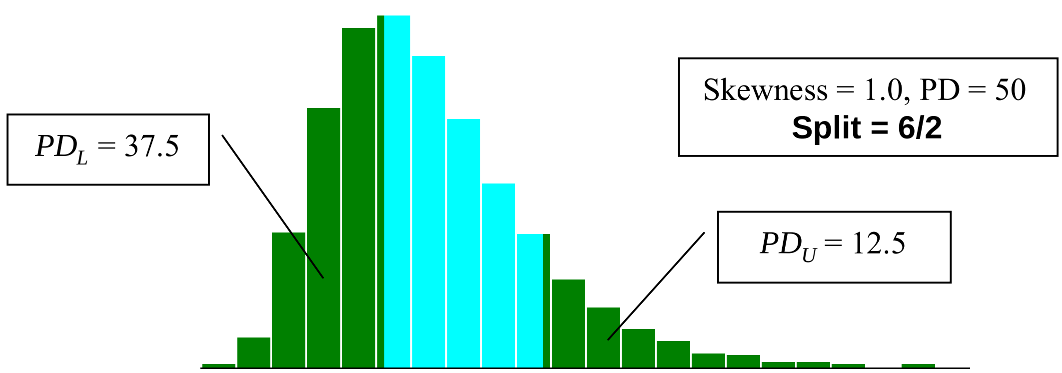


Figure 23c. Illustration of the divisions that SKEWBIAS2H uses to calculate bias in the PWL estimate for two-sided specification limits (skewness coefficient = +1.0, split=6/2).

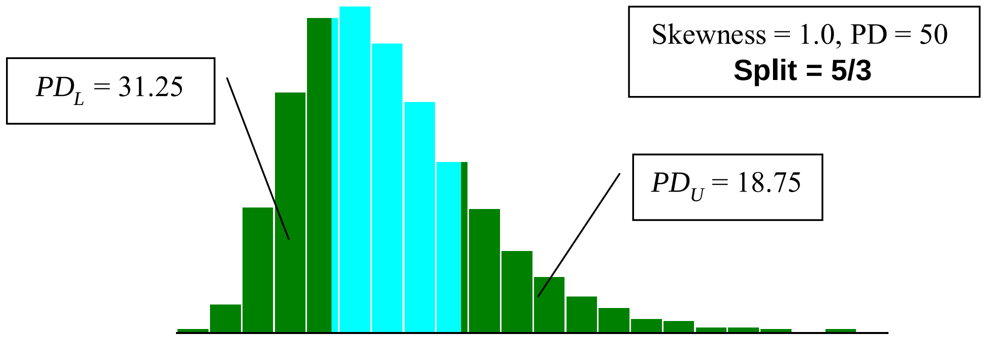


Figure 23d. Illustration of the divisions that SKEWBIAS2H uses to calculate bias in the PWL estimate for two-sided specification limits (skewness coefficient = +1.0, split=5/3).

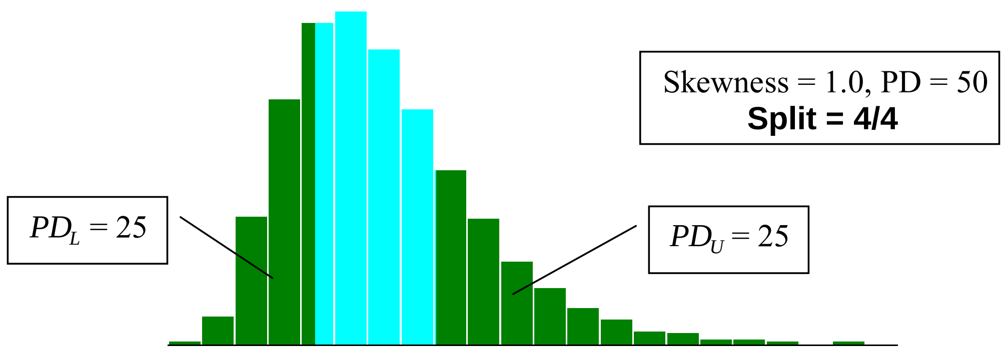


Figure 23e. Illustration of the divisions that SKEWBIAS2H uses to calculate bias in the PWL estimate for two-sided specification limits (skewness coefficient = +1.0, split=4/4).

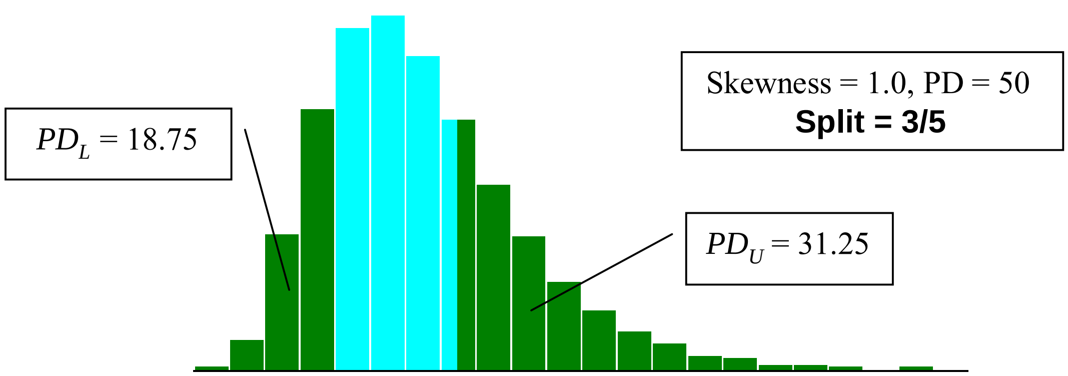


Figure 23f. Illustration of the divisions that SKEWBIAS2H uses to calculate bias in the PWL estimate for two-sided specification limits (skewness coefficient = +1.0, split=3/5).

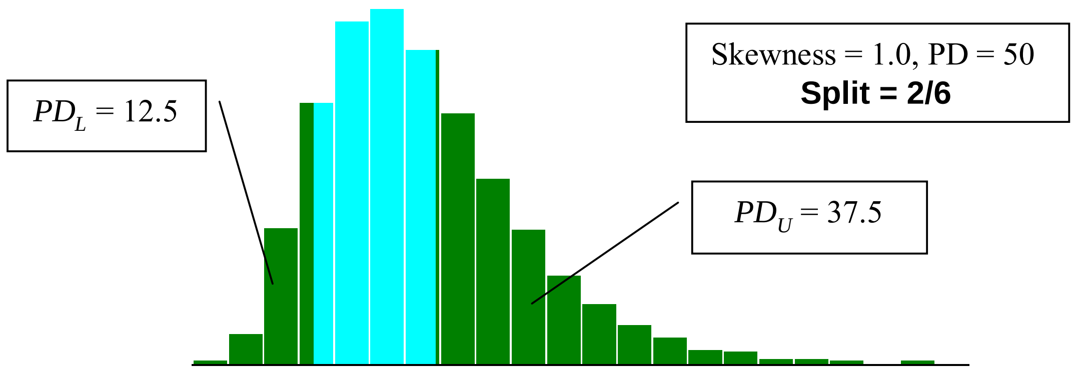


Figure 23g. Illustration of the divisions that SKEWBIAS2H uses to calculate bias in the PWL estimate for two-sided specification limits (skewness coefficient =+1.0, split = 2/6).

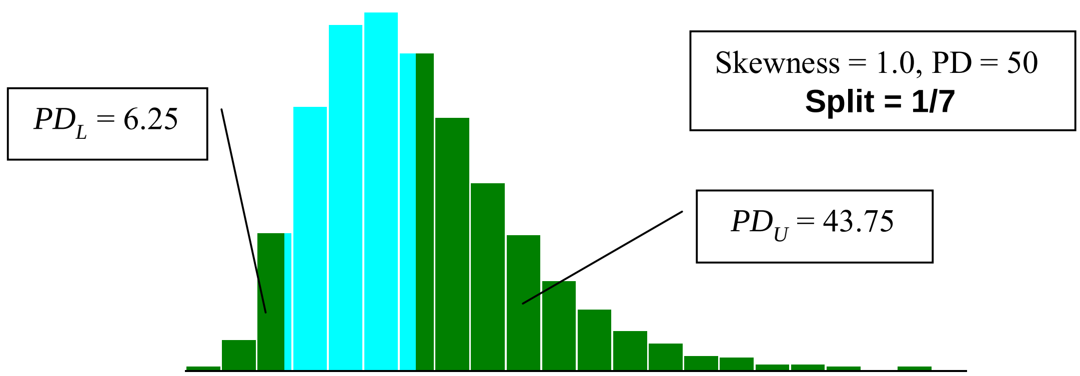


Figure 23h. Illustration of the divisions that SKEWBIAS2H uses to calculate bias in the PWL estimate for two-sided specification limits (skewness coefficient = +1.0, split = 1/7).

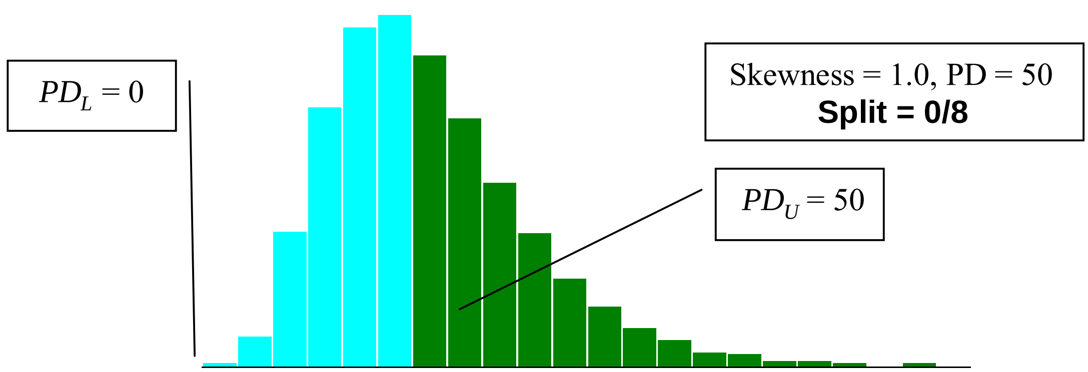


Figure 23i. Illustration of the divisions that SKEWBIAS2H uses to calculate bias in the PWL estimate for two-sided specification limits (skewness coefficient = +1.0, split = 0/8).

Tables 21 through 23 show simulation results that illustrate how the bias in estimating PD and, hence, PWL varies with the PD_L/PD_U division and sample size. In each analysis, 10,000 samples of the appropriate size were generated from a population with one of the skewness coefficients. The average bias and variability were then determined for the 10,000 PD estimates.

Figures 24 through 26 present plots of the results from tables 21 through 23, which indicate that the amount of bias in the PD estimate depends on the amount of skewness, the actual PD of the sampled population, the division of PD_L/PD_U, and the sample size used. Figures 27 through 29, which also plot the results from tables 21 through 23, highlight the effect of sample size on bias in the PD estimate.

As in the case of the one-sided specification limit shown earlier in this chapter, the larger the sample size, the greater the bias of the estimate. This probably stems from the fact that the larger sample does a better job of estimating PD for the assumed symmetrical normal distribution. This leads to a greater bias since the population is actually a skewed distribution rather than a symmetrical normal distribution.

To illustrate how this estimating bias arises for skewed distributions, figure 30 shows the plots of two populations—one that is normally distributed and one that has a skewness coefficient = +1.0. These two populations have the same mean, 5.0, and the same standard deviation, 1.0. The solid vertical line in the figure represents the mean for the two distributions. For the normal distribution, the median, or 50th percentile, occurs at the same value as the mean (i.e., 5.0). However, for the skewed distribution, the median is at a smaller value of 4.85.

Assume that the single lower specification limit is 5.0, meaning that the normal distribution has 50 PD (or 50 PWL). However, since the skewed distribution has a median of 4.85, it will have greater than 50 PD (or less than 50 PWL). Since the quality index is used to estimate PD (or PWL) and since this requires an assumption that the population is normally distributed, when the quality index determines an estimated PD of 50 for the normal distribution, the skewed distribution will have greater than 50 PD. Thus, for a population with 50 PD and for a single lower specification limit, if the skewness is away from the limit, the quality index will, on average, underestimate the PD value of the skewed distribution. If the skewness were in the direction of the lower specification limit, then the quality index would, on average, overestimate the PD value. This is exactly what is shown in figures 26 and 29. Appendix F shows additional simple illustrations of how biased estimates can be obtained.

Table 21. Results of simulations with PD = 10, 10,000 simulated lots, sample sizes = 3, 5, and 10, and two-sided limits.

PD_L/PD_U	Skewness						
	0.0	0.5	1.0	1.5	2.0	2.5	3.0
Sample Size = 3							
8/0	–0.22	+0.05	+0.73	+1.83	+2.14	+3.20	+3.80
7/1	+0.27	+0.14	+0.63	+1.34	+2.13	+3.24	+3.87
6/2	–0.14	+0.01	+1.03	+1.25	+2.61	+3.33	+4.09
5/3	–0.05	+0.12	+0.77	+1.89	+2.67	+3.87	+4.74
4/4	–0.07	+0.15	+1.07	+1.64	+3.58	+4.25	+5.43
3/5	+0.02	+0.14	+0.86	+2.17	+2.97	+4.52	+5.54
2/6	–0.15	–0.03	+1.13	+2.03	+3.58	+4.79	+6.59
1/7	–0.01	+0.43	+0.89	+1.94	+3.68	+5.31	+6.51
0/8	–0.13	–0.06	–0.13	–0.05	+0.40	+0.11	+0.31
Sample Size = 5							
8/0	+0.07	+0.71	+1.56	+2.70	+3.63	+4.58	+5.64
7/1	–0.02	+0.45	+1.44	+2.32	+3.43	+4.81	+5.78
6/2	–0.29	+0.20	+1.41	+2.46	+3.81	+4.19	+6.39
5/3	–0.23	+0.21	+1.48	+2.68	+4.47	+5.74	+6.82
4/4	–0.02	+0.36	+1.76	+3.27	+4.66	+6.20	+7.77
3/5	–0.36	+0.27	+1.80	+3.43	+5.47	+7.14	+8.28
2/6	+0.13	+0.50	+1.79	+3.79	+5.61	+7.75	+9.76
1/7	–0.01	+0.59	+1.98	+3.47	+6.19	+8.46	+10.03
0/8	+0.11	–0.26	–0.36	–0.18	+0.45	+0.45	+0.64
Sample Size = 10							
8/0	+0.05	+0.92	+2.22	+3.57	+5.07	+6.27	+7.40
7/1	+0.06	+0.56	+1.76	+3.26	+4.94	+6.26	+7.58
6/2	+0.11	+0.42	+1.66	+3.31	+5.03	+6.65	+8.25
5/3	–0.16	+0.46	+1.87	+3.64	+5.69	+7.52	+8.90
4/4	+0.07	+0.50	+2.10	+3.97	+6.13	+8.03	+9.77
3/5	+0.07	+0.60	+2.24	+4.40	+6.53	+8.79	+10.93
2/6	–0.08	+0.68	+2.37	+4.96	+7.27	+9.77	+11.91
1/7	+0.14	+0.53	+2.46	+5.32	+7.88	+10.77	+12.77
0/8	+0.02	–0.61	–0.69	–0.21	–0.10	+0.43	+1.18

Table 22. Results of simulations with PD = 30, 10,000 simulated lots, sample sizes = 3, 5, and 10, and two-sided limits.

PD_L/PD_U	Skewness						
	0.0	**0.5**	**1.0**	**1.5**	**2.0**	**2.5**	**3.0**
Sample Size = 3							
8/0	+0.11	–0.49	–0.86	–1.34	–1.46	–1.33	–1.36
7/1	–0.08	–0.64	–0.61	–1.01	–1.01	–1.13	–1.37
6/2	+0.18	–0.36	–0.68	–0.48	–0.53	–0.01	+0.12
5/3	–0.28	+0.13	+0.14	+0.40	+0.60	+1.01	+1.94
4/4	+0.34	–0.37	+0.81	+1.37	+1.73	+2.30	+3.67
3/5	–0.04	+0.63	+1.07	+2.15	+2.73	+3.75	+4.46
2/6	–0.40	+0.64	+1.86	+3.05	+4.18	+5.50	+6.67
1/7	–0.47	+0.59	+2.12	+3.66	+5.46	+7.58	+8.19
0/8	+0.03	+0.27	+0.78	+1.19	+1.84	+2.10	+2.67
Sample Size = 5							
8/0	–0.07	–0.84	–1.96	–2.26	–2.69	–2.94	–3.08
7/1	+0.35	–1.36	–2.02	–2.46	–2.52	–2.61	–2.32
6/2	+0.13	–0.75	–1.31	–1.27	–1.06	–0.56	–0.27
5/3	+0.20	–0.50	–0.66	–0.01	+1.04	+1.48	+1.80
4/4	+0.04	+0.18	+0.88	+2.04	+2.75	+4.08	+5.25
3/5	–0.02	+0.61	+1.97	+3.61	+5.10	+6.63	+7.79
2/6	–0.26	+1.17	+3.19	+5.18	+7.25	+9.28	+11.22
1/7	+0.02	+1.87	+3.99	+7.05	+9.18	+12.10	+14.07
0/8	–0.03	+1.18	+1.81	+3.69	+4.31	+5.27	+5.99
Sample Size = 10							
8/0	+0.03	–1.44	–2.40	–3.14	–3.56	–3.63	–3.53
7/1	+0.12	–2.01	–2.97	–3.34	–3.36	–3.04	–2.80
6/2	+0.07	–1.28	–2.21	–1.83	–1.24	–0.57	+0.45
5/3	–0.10	–0.75	–0.43	+0.36	+1.36	+2.59	+3.83
4/4	–0.01	+0.33	+1.28	+2.57	+4.55	+6.19	+8.34
3/5	+0.06	+1.11	+3.08	+5.29	+7.73	+10.08	+12.51
2/6	–0.23	+2.13	+5.13	+7.98	+10.88	+13.38	+15.99
1/7	–0.01	+2.59	+6.44	+10.01	+13.97	+16.84	+19.92
0/8	+0.16	+1.56	+3.35	+5.40	+6.80	+8.17	+9.70

Table 23. Results of simulations with PD = 50, 10,000 simulated lots, sample sizes = 3, 5, and 10, and two-sided limits.

PD_L/PD_U	Skewness						
	0.0	0.5	1.0	1.5	2.0	2.5	3.0
Sample Size = 3							
8/0	–0.08	–0.82	–1.37	–2.29	–3.12	–3.38	–3.40
7/1	–0.39	–0.94	–1.77	–2.11	–2.39	–3.13	–3.11
6/2	+0.41	–0.67	–1.42	–1.64	–1.59	–2.45	–2.33
5/3	+0.05	–0.48	–0.40	–0.77	–0.48	–0.60	+0.08
4/4	+0.13	–0.15	+0.59	+0.55	+1.00	+1.47	+1.96
3/5	+0.25	+0.44	+0.97	+1.97	+2.56	+2.97	+4.16
2/6	–0.04	+1.12	+1.69	+2.91	+4.14	+5.32	+5.71
1/7	–0.12	+1.38	+2.64	+3.86	+5.23	+6.80	+8.22
0/8	+0.05	+0.87	+1.41	+1.99	+2.92	+3.03	+4.11
Sample Size = 5							
8/0	–0.15	–1.44	–3.06	–4.67	–5.64	–6.45	–7.55
7/1	–0.01	–2.15	–3.56	–4.90	–5.80	–6.24	–6.57
6/2	+0.06	–1.39	–2.70	–3.46	–3.79	–3.97	–3.97
5/3	+0.25	–1.16	–1.07	–1.11	–0.86	–0.60	–0.34
4/4	–0.19	–0.04	+0.48	+1.17	+2.28	+2.91	+3.74
3/5	+0.03	+0.89	+2.22	+3.53	+5.00	+6.51	+7.51
2/6	–0.02	+1.73	+3.88	+6.27	+8.03	+9.95	+11.50
1/7	0.00	+2.41	+5.30	+8.06	+10.93	+12.89	+15.06
0/8	+0.07	+1.65	+3.17	+4.57	+5.42	+6.41	+7.43
Sample Size = 10							
8/0	–0.07	–2.26	–4.66	–6.69	–8.06	–9.33	–9.98
7/1	–0.17	–3.22	–5.56	–6.96	–8.20	–8.83	–9.08
6/2	+0.15	–2.22	–3.78	–4.69	–4.71	–4.90	–4.67
5/3	+0.26	–1.01	–1.56	–1.31	–0.48	+0.41	+0.94
4/4	+0.17	+0.04	+1.15	+2.38	+4.09	+5.65	+6.74
3/5	–0.01	+1.41	+3.70	+5.56	+8.17	+10.36	+12.33
2/6	0.00	+2.62	+5.93	+8.88	+11.84	+14.59	+17.13
1/7	–0.05	+3.48	+7.56	+11.71	+15.36	+18.38	+20.71
0/8	–0.08	+2.34	+4.93	+6.61	+8.03	+9.35	+10.39

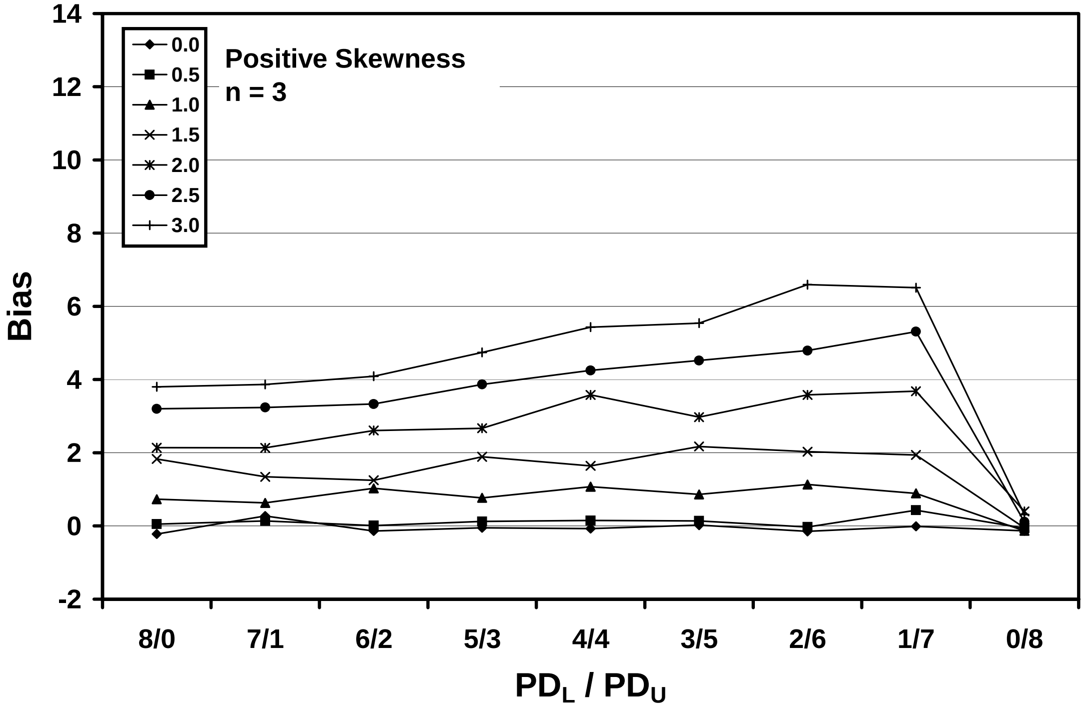


Figure 24a. Plot of bias versus PD_L/PD_U divisions for 10,000 simulated lots with PD = 10, sample = 3, and two-sided limits.

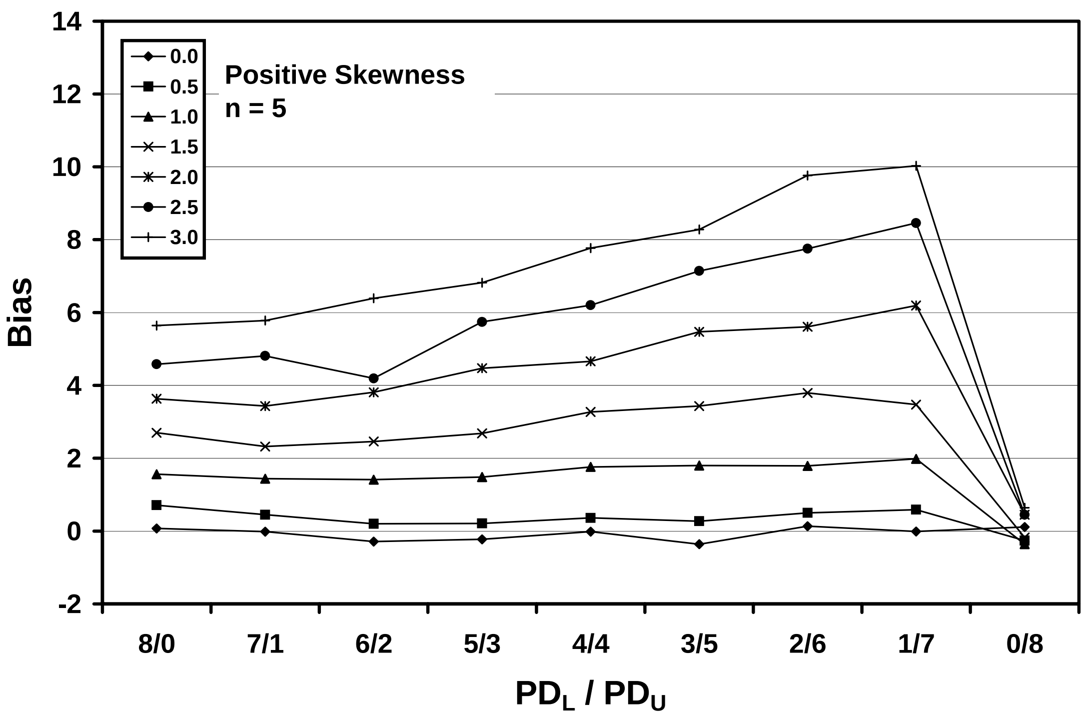


Figure 24b. Plot of bias versus PD_L/PD_U divisions for 10,000 simulated lots with PD = 10, sample = 5 and two-sided limits.

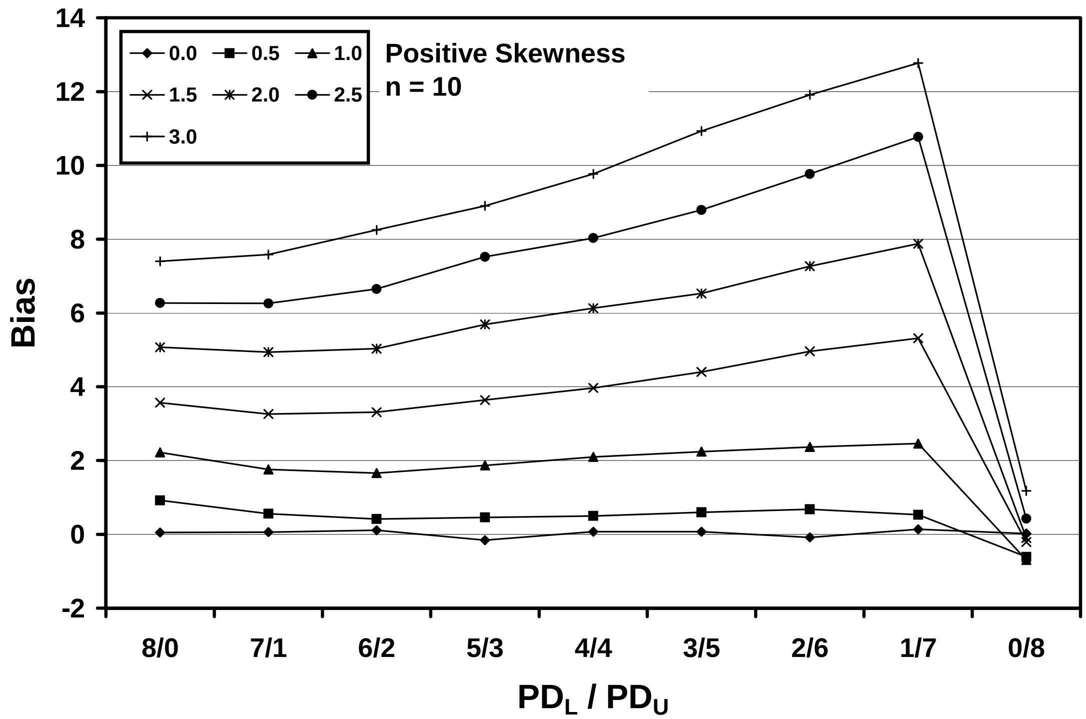


Figure 24c. Plot of bias versus PD_L/PD_U divisions for 10,000 simulated lots with PD = 10, sample = 10, and two-sided limits.

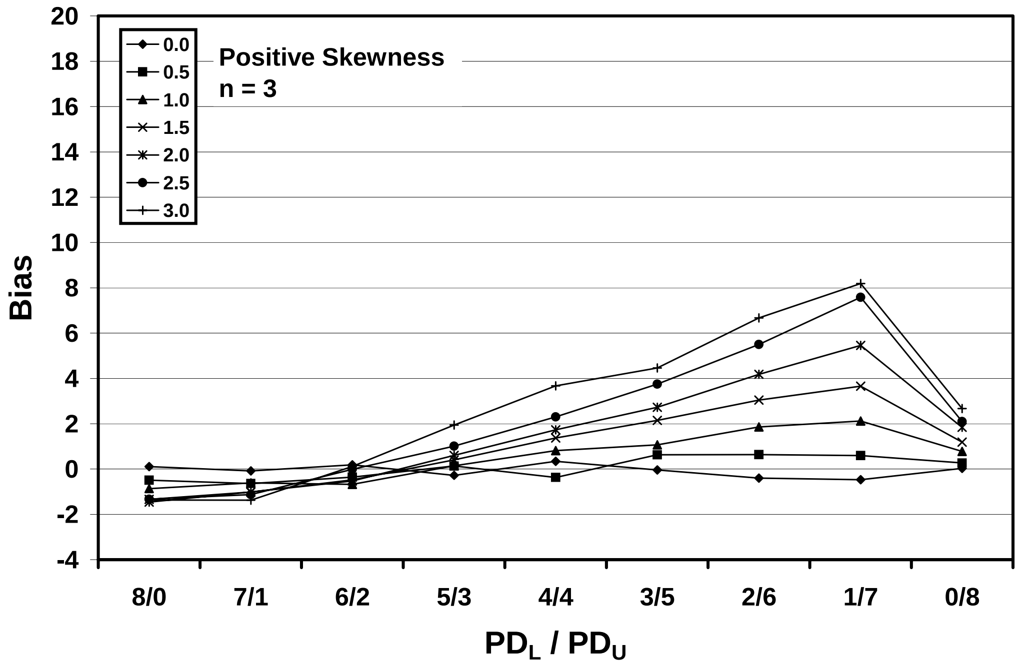


Figure 25a. Plot of bias versus PD_L/PD_U divisions for 10,000 simulated lots with PD = 30, sample = 3, and two-sided limits.

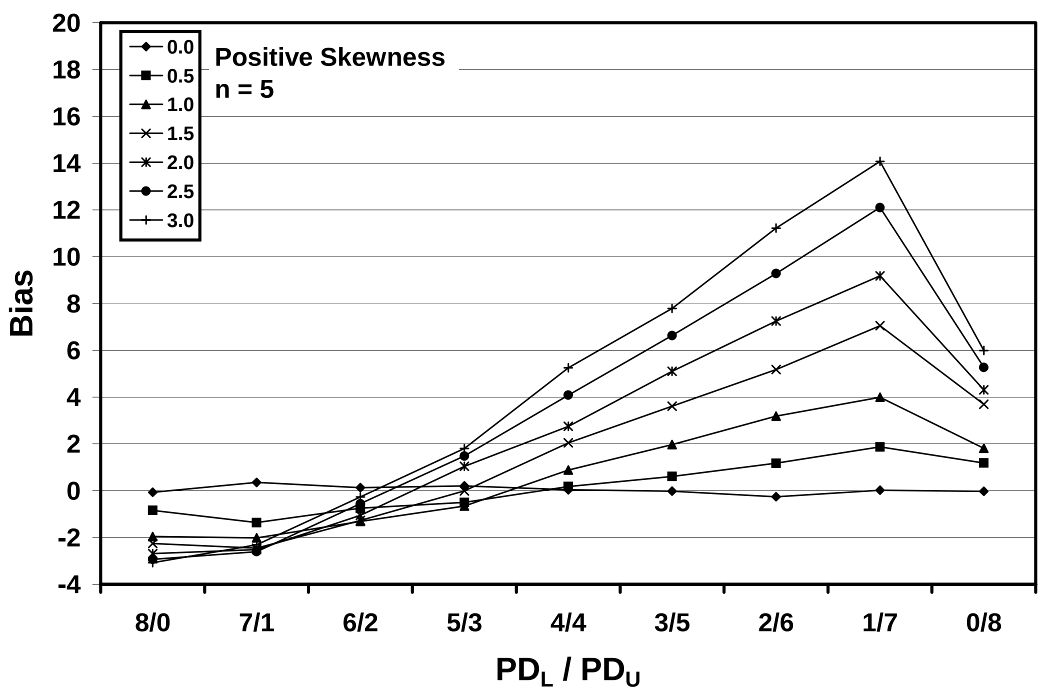


Figure 25b. Plot of bias versus PD_L/PD_U divisions for 10,000 simulated lots with PD = 30, sample = 5, and two-sided limits.

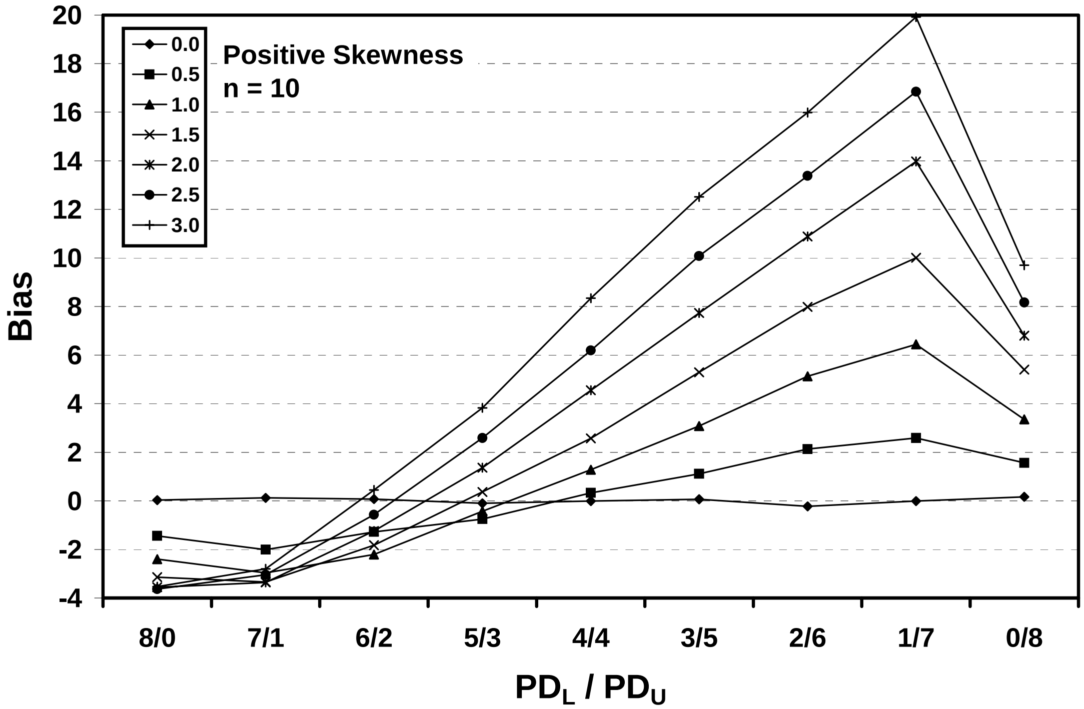


Figure 25c. Plot of bias versus PD_L/PD_U divisions for 10,000 simulated lots with PD = 30, sample = 10, and two-sided limits.

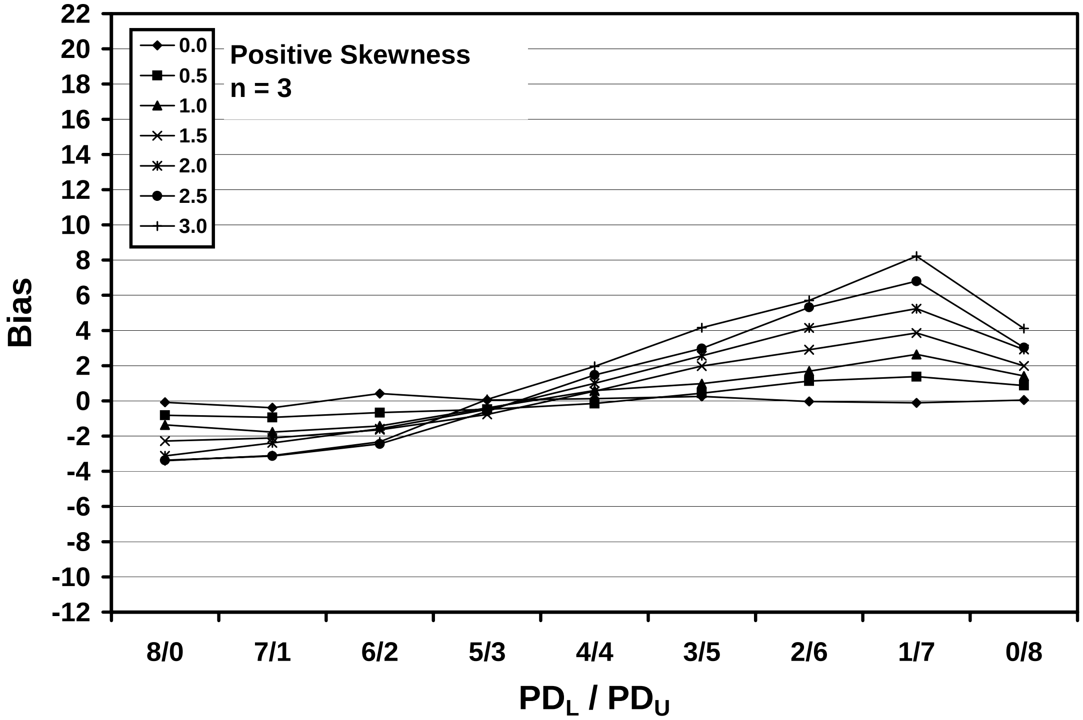


Figure 26a. Plot of bias versus PD_L/PD_U divisions for 10,000 simulated lots with PD = 50, sample = 3 and two-sided limits.

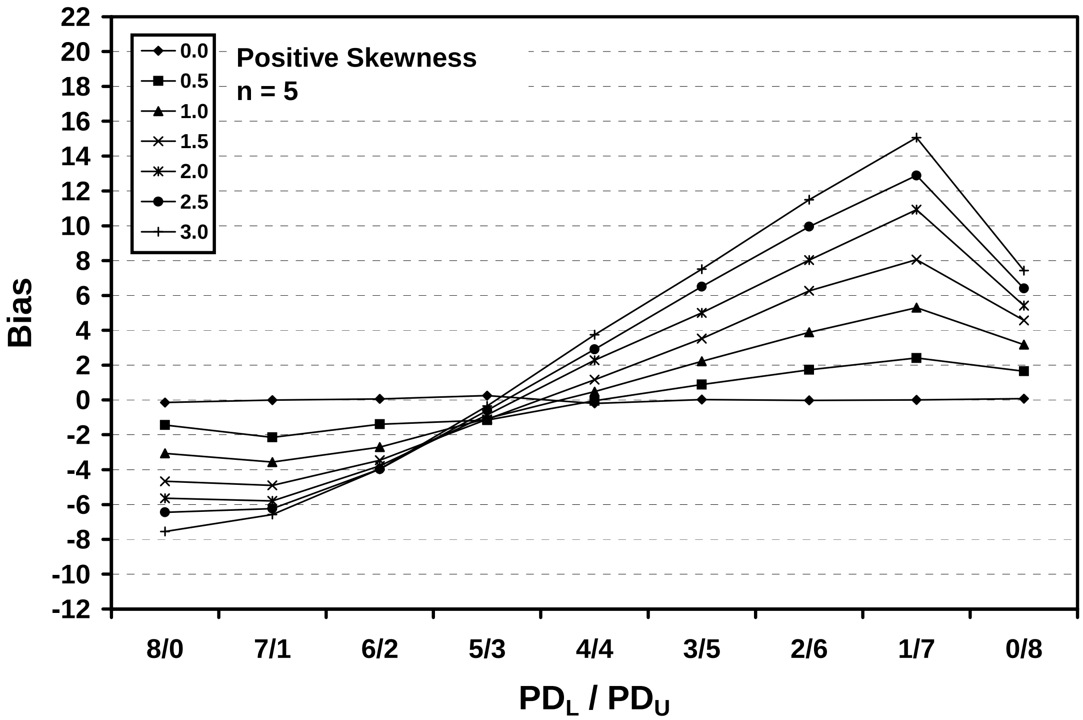


Figure 26b. Plot of bias versus PD_L/PD_U divisions for 10,000 simulated lots with PD = 50, sample = 5, and two-sided limits.

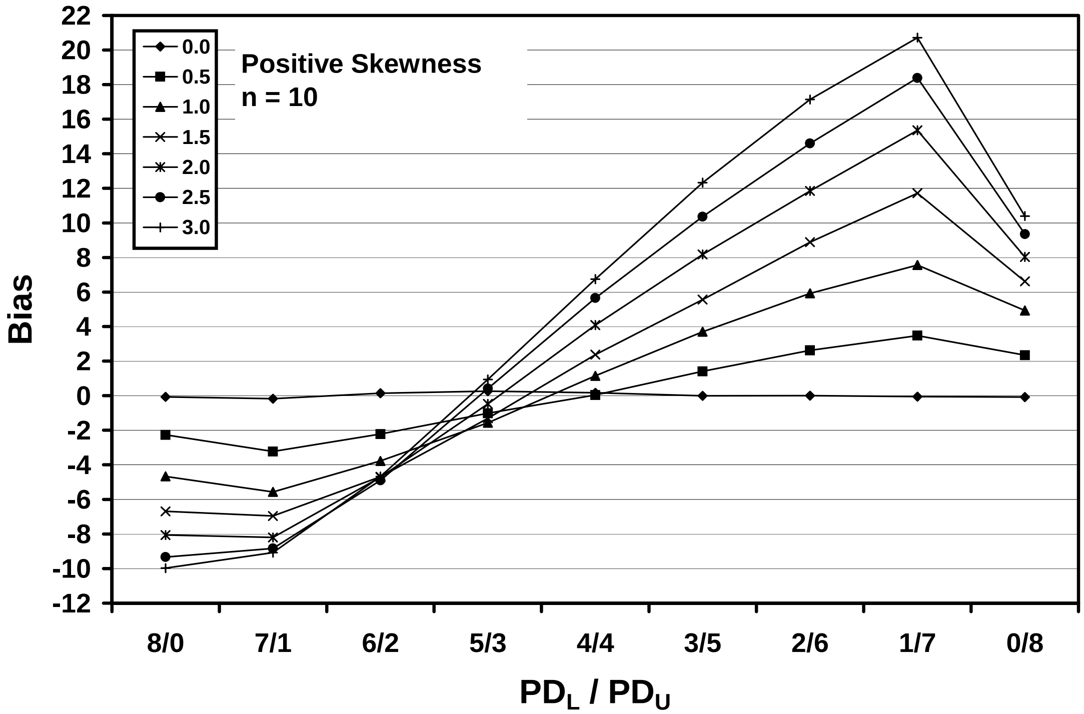


Figure 26c. Plot of bias versus PD_L/PD_U divisions for 10,000 simulated lots with PD = 50, sample = 10, and two-sided limits.

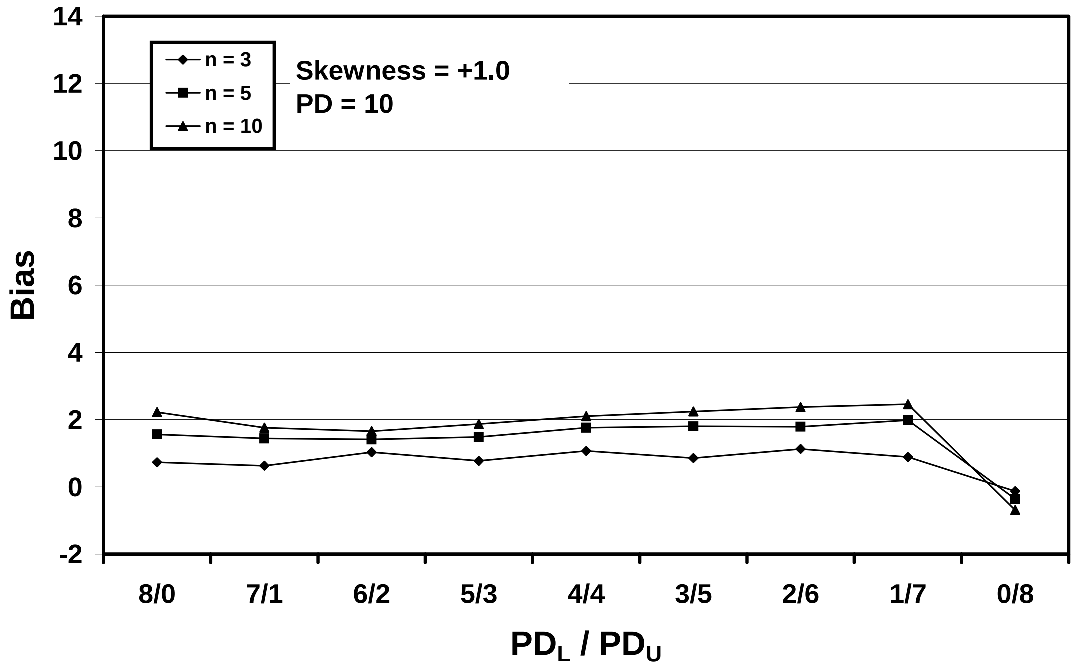


Figure 27a. Plot of bias versus SKEWBIAS2H divisions for 10,000 simulated lots with PD = 10, skewness = 1, and two-sided limits.

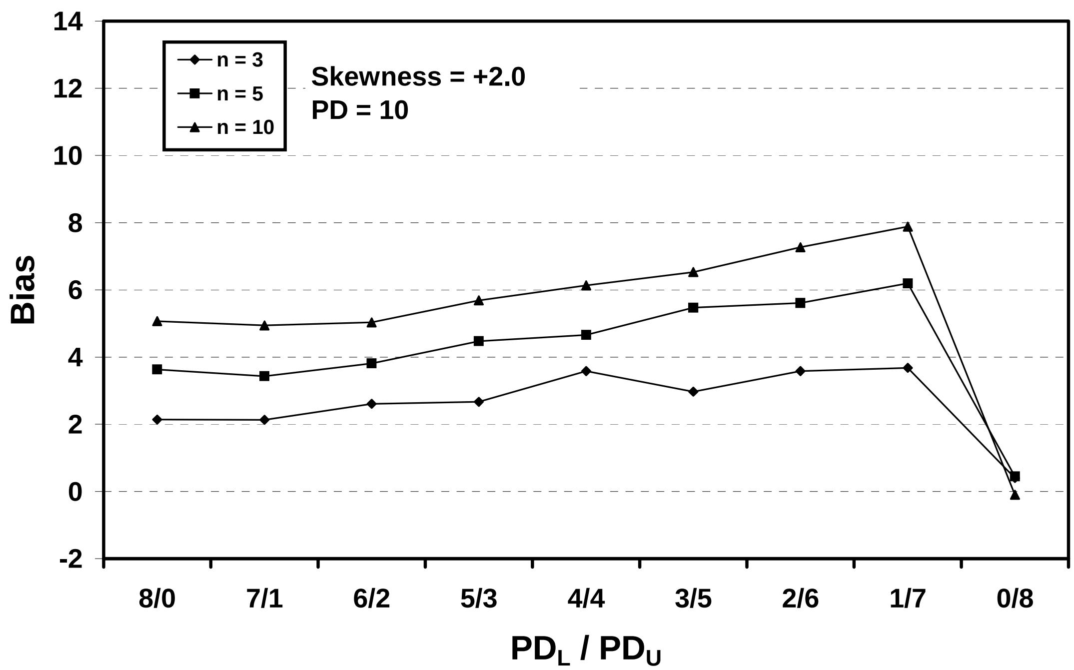


Figure 27b. Plot of bias versus SKEWBIAS2H divisions for 10,000 simulated lots with PD = 10, skewness = 2, and two-sided limits.

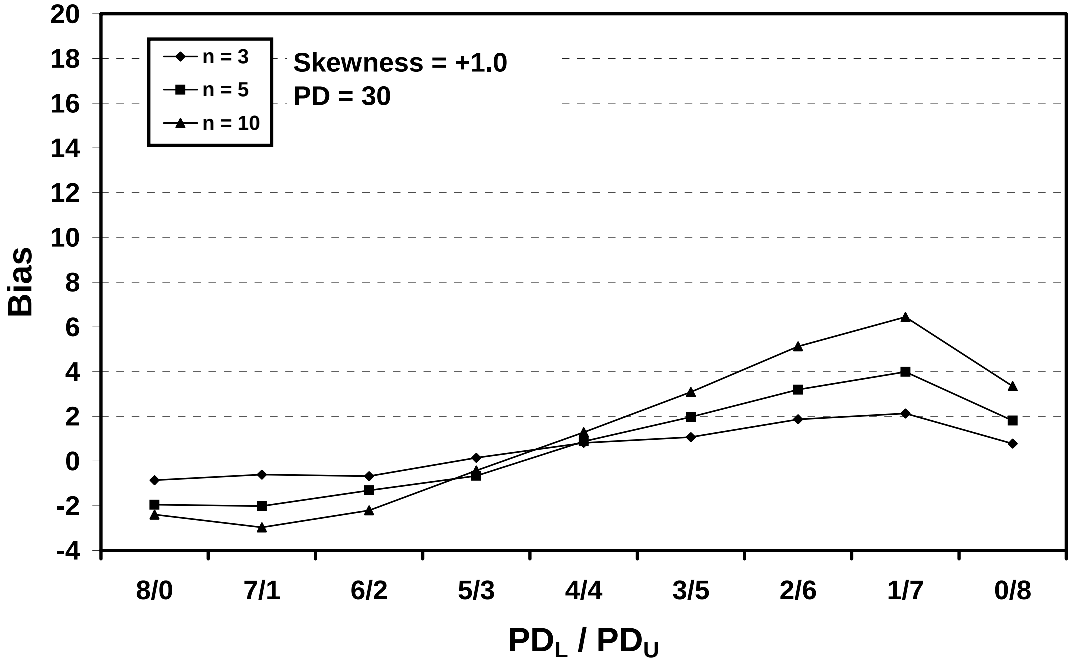


Figure 28a. Plot of bias versus SKEWBIAS2H divisions for 10,000 simulated lots with PD = 30, skewness = 1, and two-sided limits.

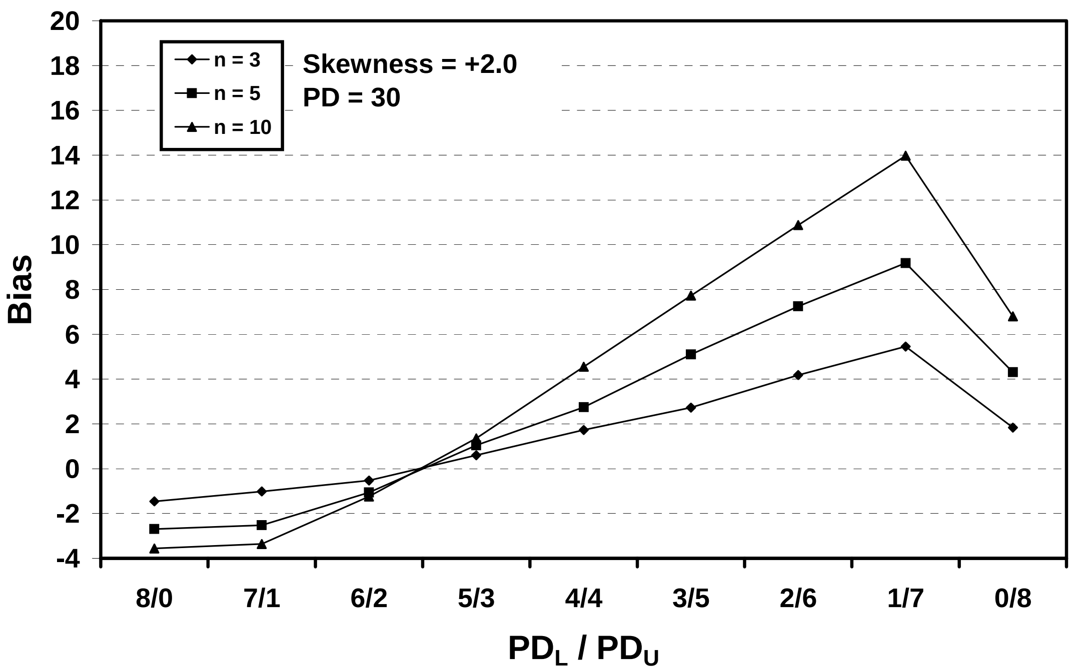


Figure 28b. Plot of bias versus SKEWBIAS2H divisions for 10,000 simulated lots with PD = 30, skewness = 2, and two-sided limits.

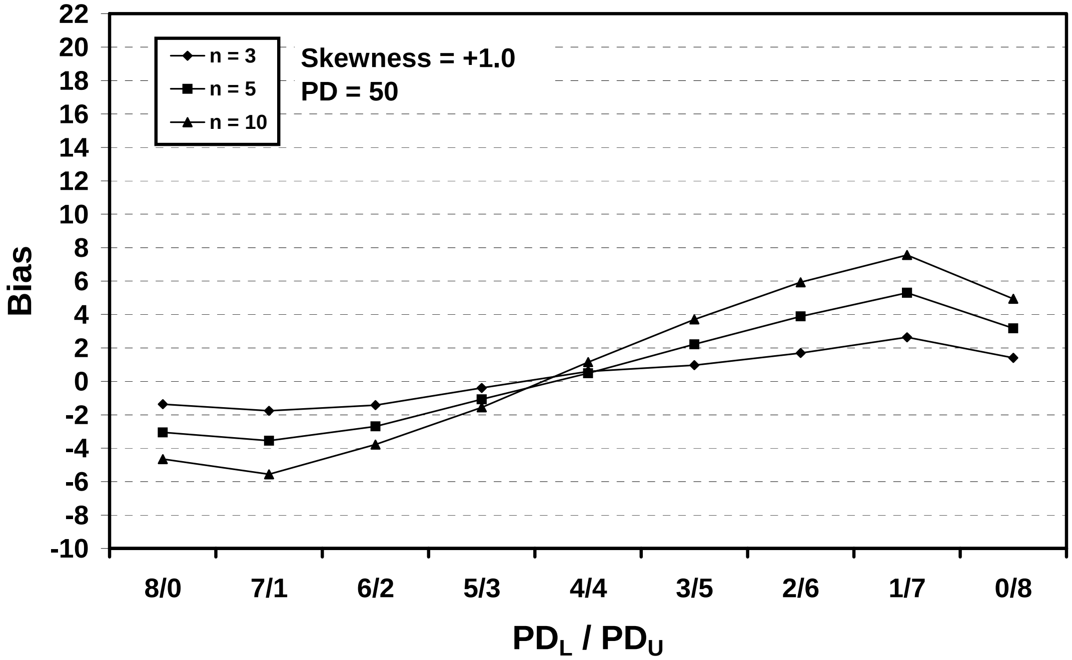


Figure 29a. Plot of bias versus SKEWBIAS2H divisions for 10,000 simulated lots with PD = 50, skewness = 1, and two-sided limits.

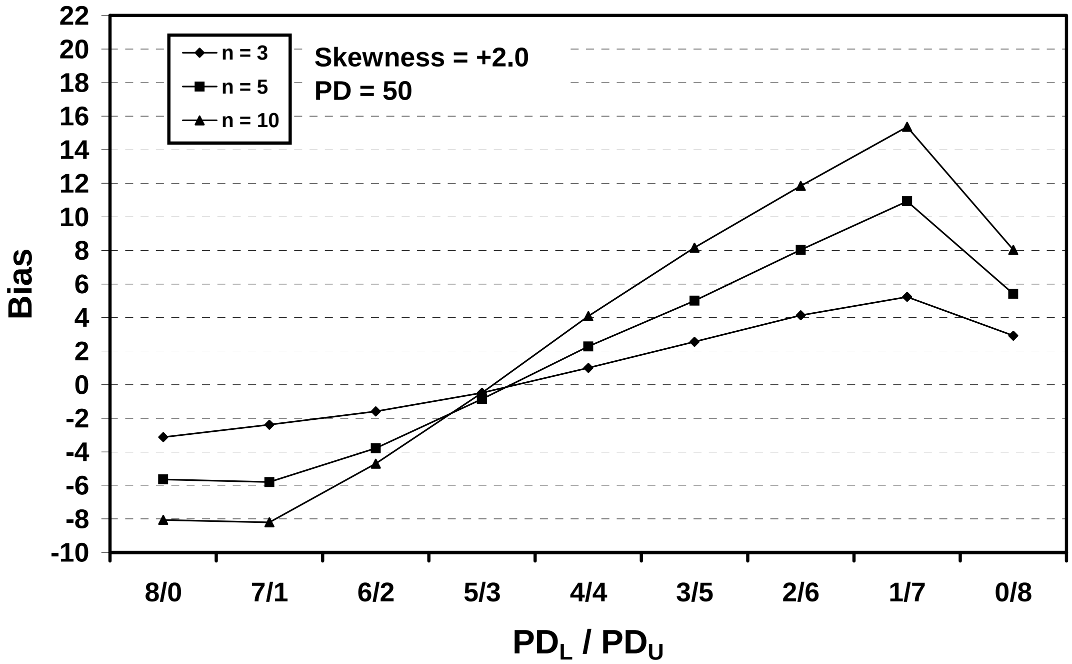


Figure 29b. Plot of bias versus SKEWBIAS2H divisions for 10,000 simulated lots with PD = 50, skewness = 2, and two-sided limits.

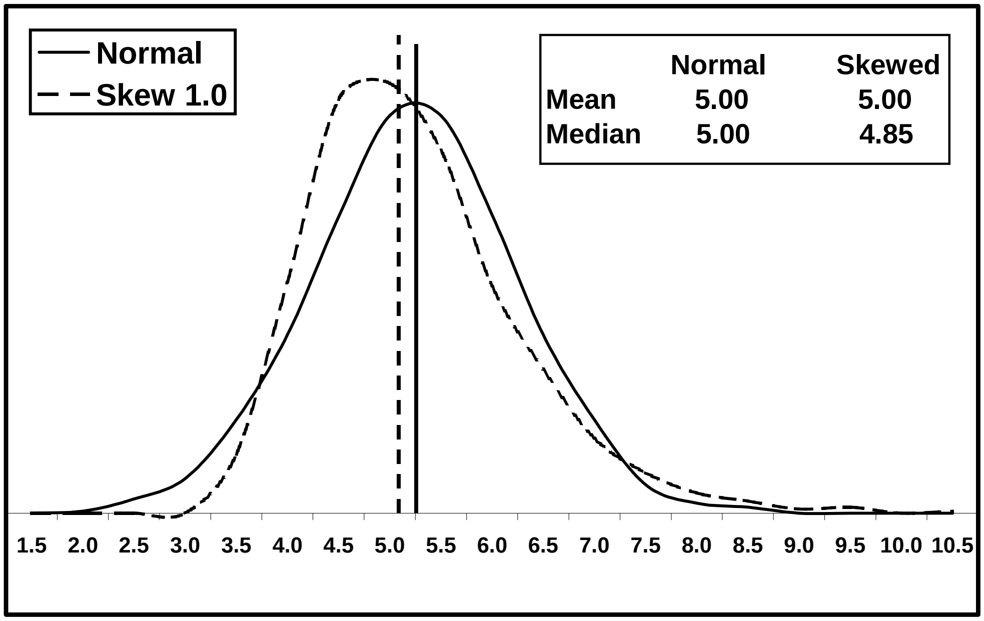


Figure 30. Comparison of a normal population with a population with skewness coefficient = +1.0.

Variability of PWL/PD Estimates: Simulation was also used to evaluate the variability of PD estimates and to show how sample size affects this variability. Figure 31 shows an example of an output screen for a normal population (i.e., skewness coefficient = 0.00) that actually has a PD = 10 and sample size = 5. The results of the nine divisions for PD_L/PD_U are clearly indicated.

The BIAS/SE results indicate the average bias and the standard error of the bias, respectively. The bias values are relatively closely distributed about zero, indicating that the method for estimating PD (and, hence, PWL) is not biased for a symmetrical distribution. This is to be expected since the PD estimating method assumes a symmetrical normal distribution.

The standard error values for sample size $n = 5$ are all 0.11. The standard error is determined by calculating the standard deviation of the 10,000 lot sample means and then dividing this standard deviation by the square root of 10,000 (i.e., the number of sample averages). The standard error is a measure of the variability of the estimate for the population mean. The standard error is not a measure of the variability of the individual lot sample means. However, it can be related to this individual sample mean variability.

Table 8 shows the standard deviation values calculated from 1000 sample means. If the standard deviation values in table 8 were taken as estimates for the variability of the individual sample means, then these values would estimate the standard error values in figure 31 by dividing the table 8 values by the square root of 10,000, which is 100. The value in table 8 for a population with 90 PWL, which corresponds to 10 PD in figure 31, is 11.42. If you divide this number by 100, the result is 0.1142, which is consistent with the standard error of 0.11 shown in figure 31.

Figure 32 shows portions of the output from three simulations, each with a population of PD = 10 and skewness coefficient = 0.00. In the figure, the divisions for 8/0, 7/1, and 6/2 are shown for sample sizes = 3, 5, and 10. These histograms represent the individual bias values for the 10,000 lots that were simulated.

One thing is immediately apparent—the bias plots are not symmetrical. This is because of the natural boundaries of 0 PD and 100 PD. It is not possible to have less than 0 PD, so it is not possible to have a bias estimate of less than –10 PD. It is possible, however, to overestimate PD by as much as +90 PD. This is reflected in the plots that are skewed to the right, while having no bias values of less than –10. The second point to notice in the plots is that the variability of the bias results is related to sample size. This is shown in the standard error values and in the spread of the histogram plots. The larger the sample size, the smaller the standard error and the less spread there is in the bias histogram.

As already shown in tables and figures above, the bias values vary with the amount of skewness, the sample size, the PD or PWL of the population, and the division of PD material outside the upper and lower specification limits. Figure 32 shows that the shape and the spread of the distribution of the sample means vary with the sample size. These distributions also vary in shape considerably with the population PD or PWL value. Appendix G contains a number of sample output screens that illustrate the shape and spread of the sample means for a variety of sample sizes and population PD values.

While appendix G contains populations with skewness coefficients up to 3.00, in practice, it is unlikely that highway materials will have skewness values much greater than 1.00. Figure 33 shows an example of an output screen for a population with a skewness coefficient = 1.00 that has a PD = 30 and sample size = 5. Figure 34 shows portions of the output screens for the same population, but with sample sizes = 3, 5, and 10.

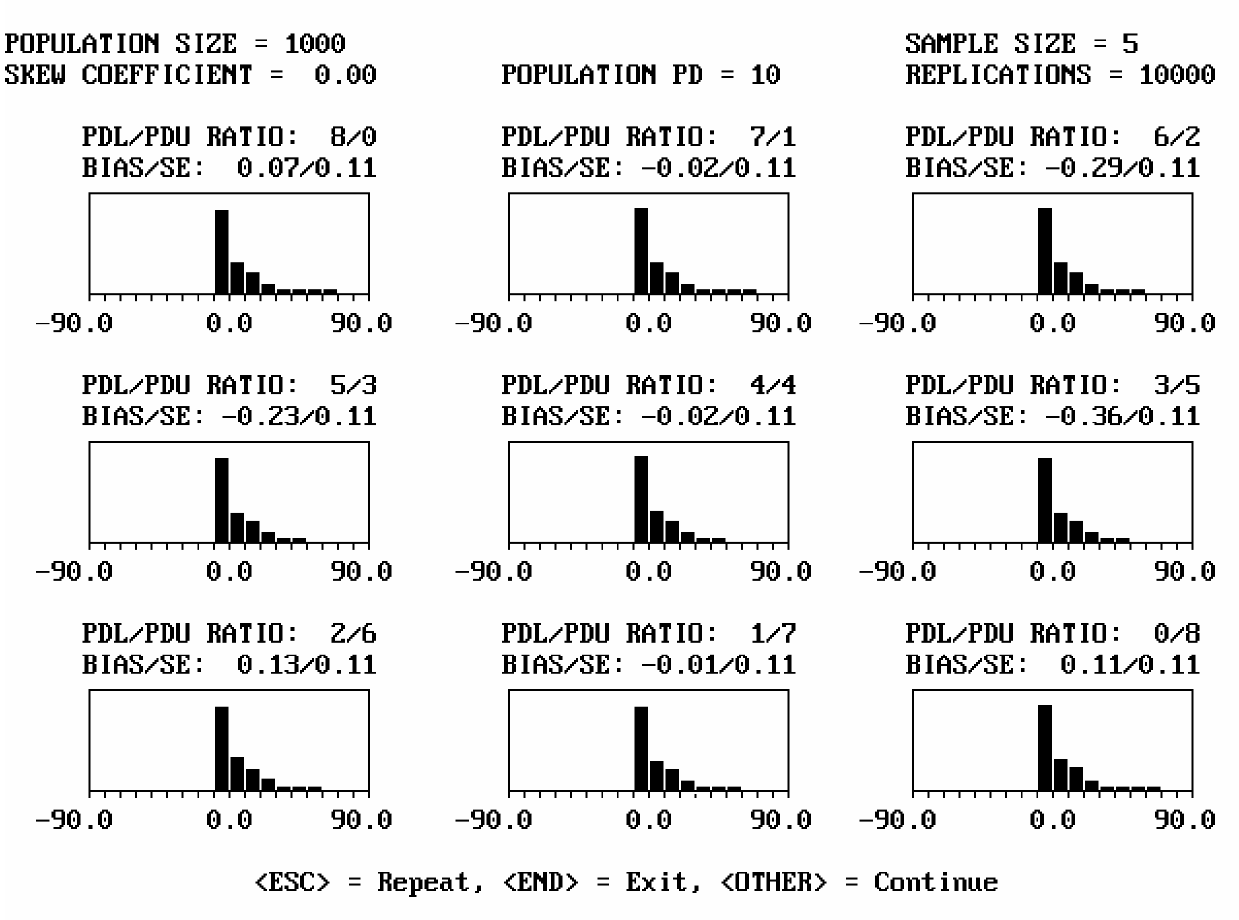


Figure 31. Sample program output screen for a population with PD = 10, skewness coefficient = 0.00, and sample size = 5.

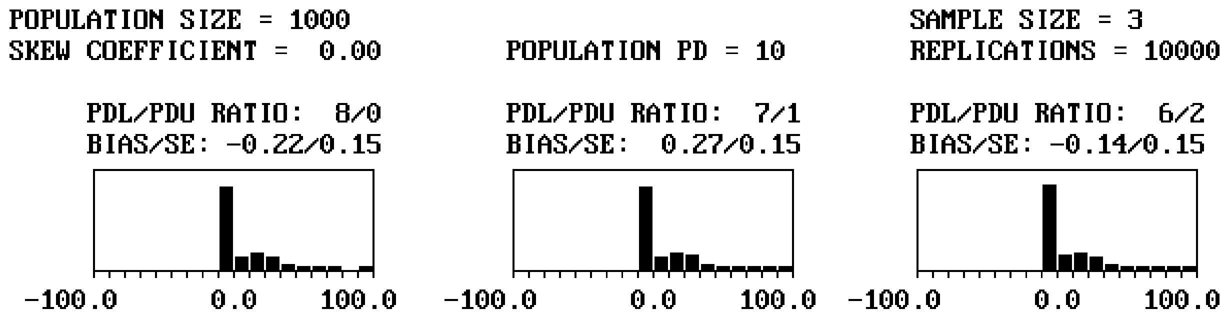


Figure 32a. Portions of program output screen for PD = 10, skewness coefficient = 0.00, sample=3.

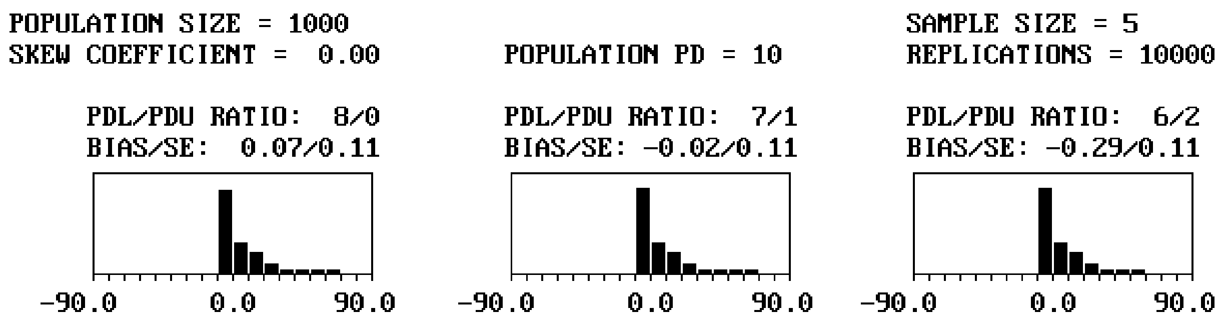


Figure 32b. Portion of program output screens for PD = 10, skewness coefficient = 0.00, sample = 5.

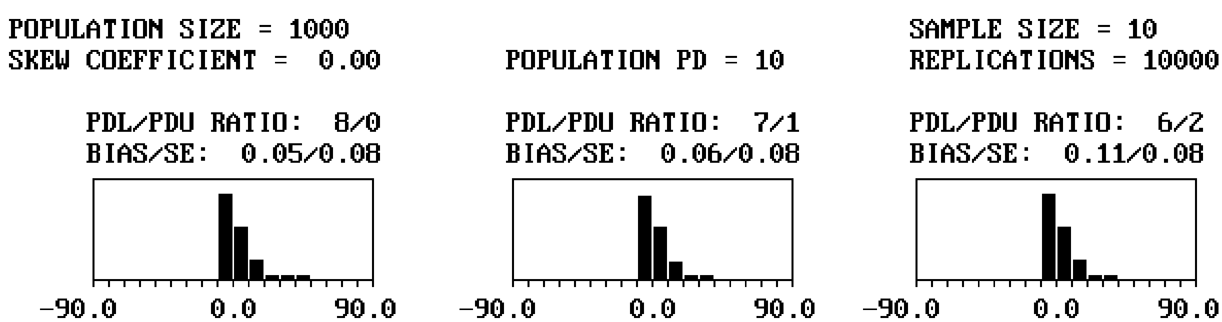


Figure 32c. Portion of program output screens for PD = 10, skewness coefficient = 0.00, sample = 10.

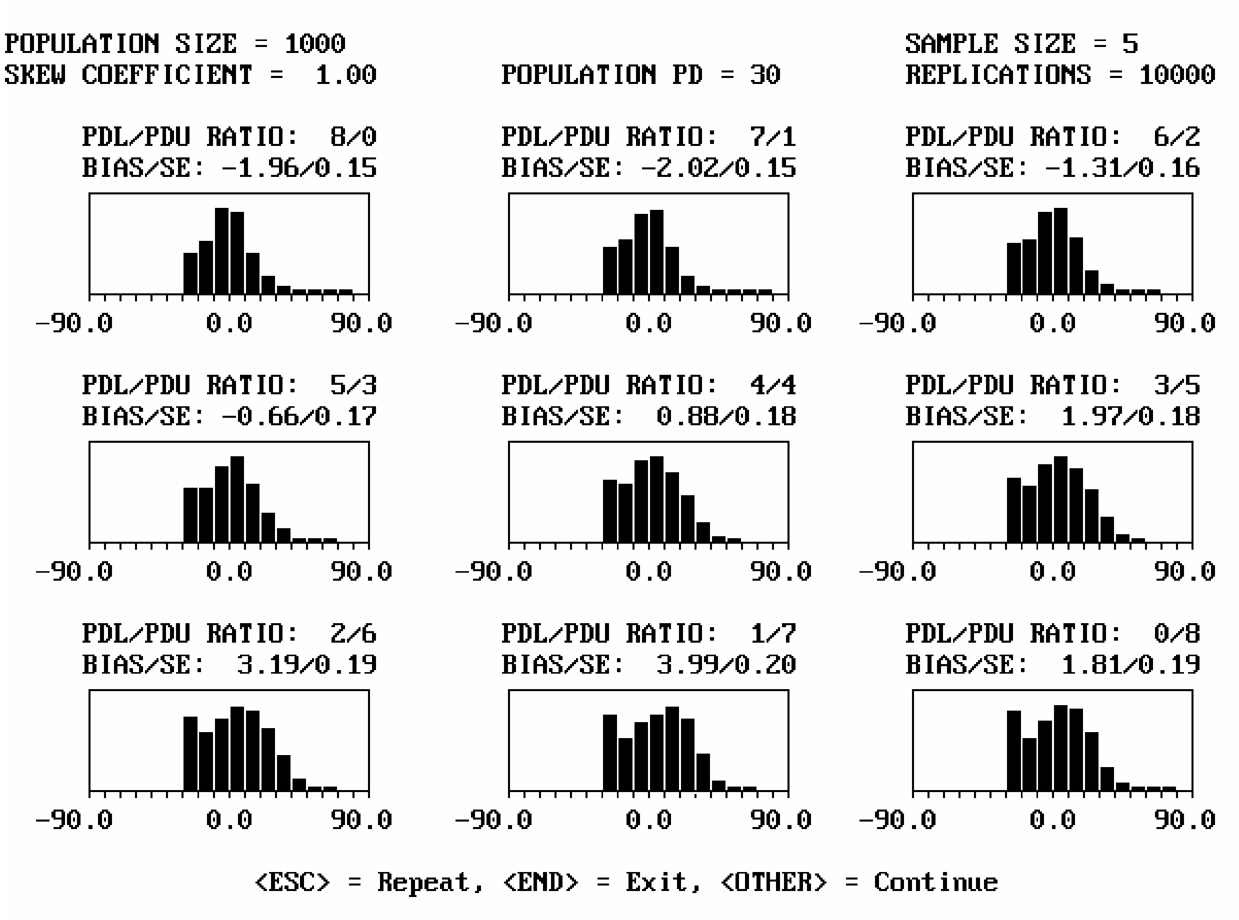


Figure 33. Sample program output screen for a population with PD = 30, skewness coefficient = 1.00, and sample size = 5.

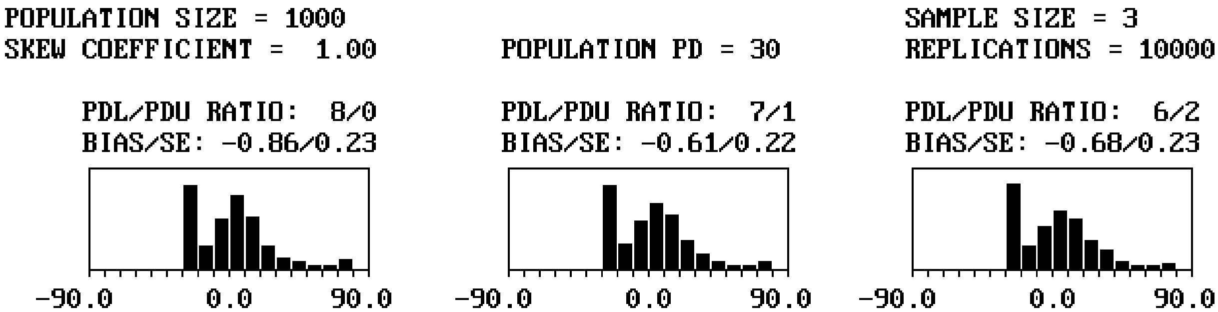


Figure 34a. Portion of program output screens for PD = 30, skewness coefficient = 1.00, and sample = 3.

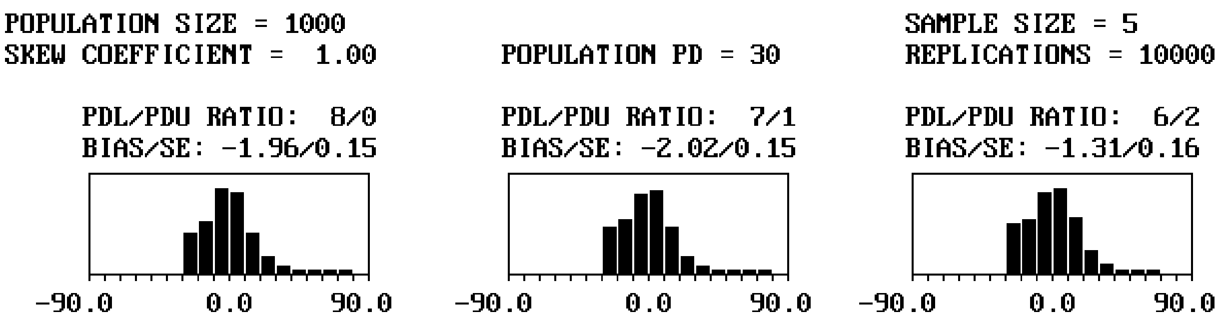


Figure 34b. Portion of program output screens for PD = 30, skewness coefficient = 1.00, and sample = 5.

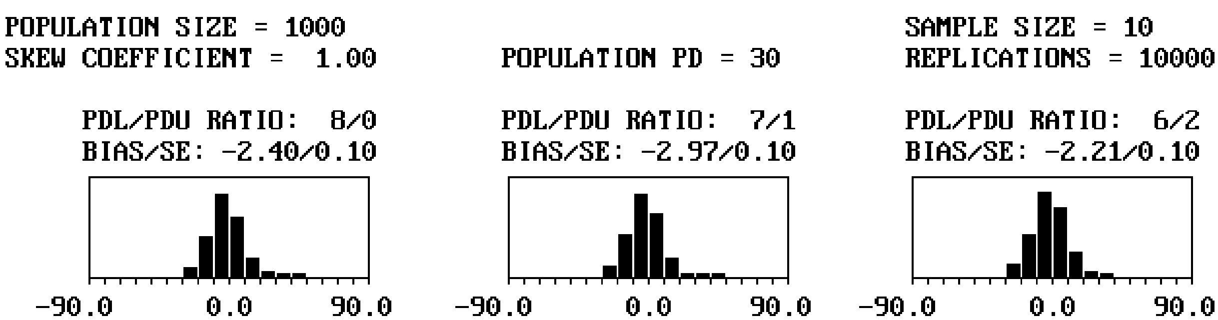


Figure 34c. Portion of program output screens for PD = 30, skewness coefficient = 1.00, and sample = 10.

AAD Evaluation

A computer simulation program was developed to investigate the performance of the AAD estimating procedures when working with a skewed distribution. The program can generate a population with any mean and standard deviation, and with skewness coefficients for distributions that are symmetrical (0.0), negatively skewed (–3.0, –2.5, –2.0, –1.5, –1.0, –0.5), or positively skewed (+3.0, +2.5, +2.0, +1.5, +1.0, +0.5). The program determines the actual AAD for the population with the input values for mean, standard deviation, skewness coefficient, and offset of the mean from the target value. The program then generates 10,000 samples from the input distribution and determines the estimated AAD for each sample. The sample size can be 3 to 30 values per sample. The program then determines the mean, standard deviation, standard error, minimum value, and maximum value for the 10,000 AAD values. It also prints a histogram for the generated AAD values.

Figure 35 shows the effect of a skewed population on the actual population AAD for populations that are centered on the target value and that have skewness coefficients ranging from 0.0 to +3.0 in increments of 0.5. The actual AAD values would be the same for negative skewness coefficients of the same magnitude as long as the population means were centered on the target. Since the population means are centered on the target values, the population AAD values actually decrease as the populations become more skewed.

Figure 36 shows the effect of sample sizes = 3, 5, and 10 on the sampling distributions for the AAD estimates for populations that are centered on the target value and that have skewness coefficients of 0.0, +0.5, and +1.0. As would be expected, the spread of the AAD estimates decreases as the sample size increases. This is reflected in the reduced spread in the histogram plots and in the smaller standard deviation values as the sample size increases. It also appears that the mean of the sampling distribution approximates the actual population AAD value quite well regardless of the level of skewness in the population.

Figure 37 shows three populations with the same standard deviation, but with mean offsets from the target = 0.0, 1.0, and 2.0. These mean offsets are measured in standard deviation units. As would be expected, the actual population AAD increases as the mean moves away from the target value. The mean of the estimated AAD values is quite close to the actual values regardless of the mean offset. However, the standard deviation (i.e., the spread) of the AAD sample values increases as the mean offset increases.

Figure 38 shows three populations that have extremely different distributions, but which all have essentially equal AAD values. Although the means of the sampling distributions are about the same, the skewed distributions have greater spread in the AAD values for individual lots. It is not certain, but it seems unlikely that each of these populations would perform identically in service, even though they have the same population AAD values.

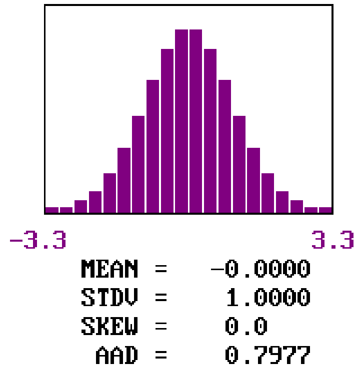


Figure 35a. Example of AAD distribution and actual values for populations centered on the target and with skewness coefficient of 0.

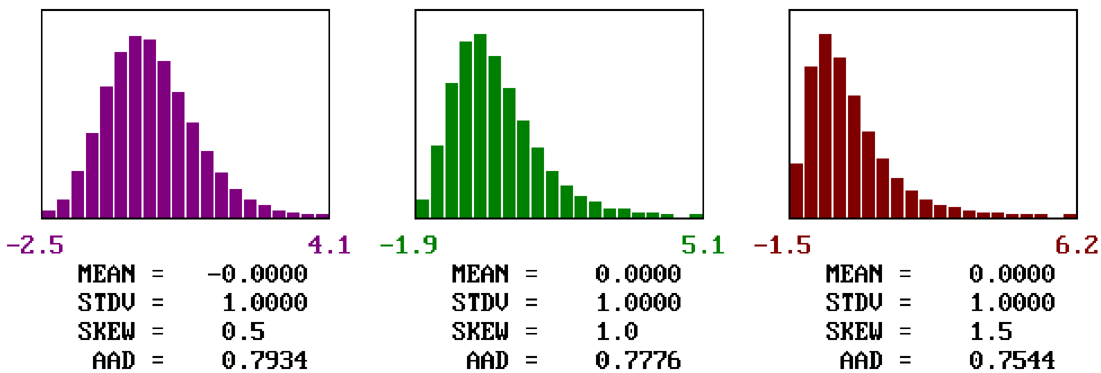


Figure 35b. Examples of AAD distribution and actual values for populations centered on the target and with skewness coefficients between .5 and 1.5.

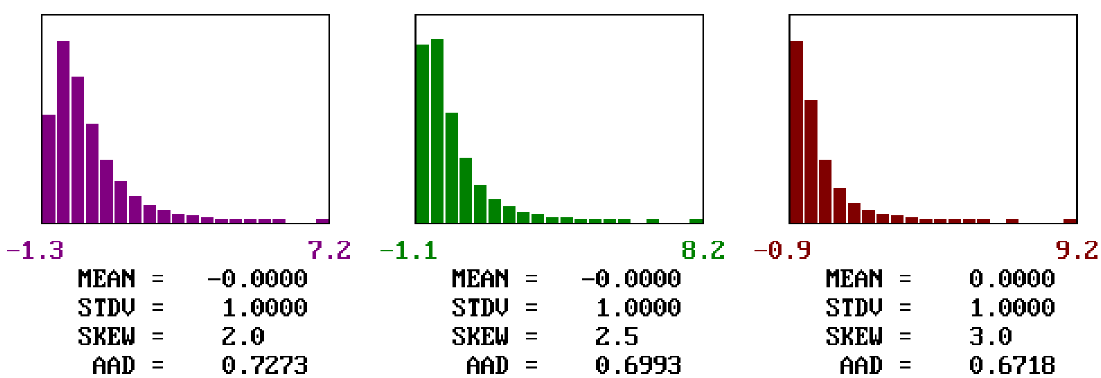


Figure 35c. Examples of AAD distribution and actual values for populations centered on the target and with skewness coefficients between 2.0 and 3.0.

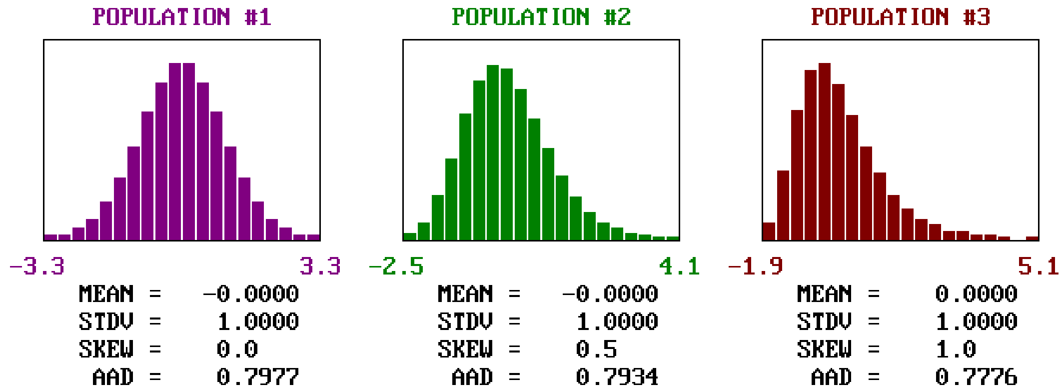


Figure 36a. Comparison of the shapes and spread of estimated AAD values for populations centered on the target and with various skewness coefficients.

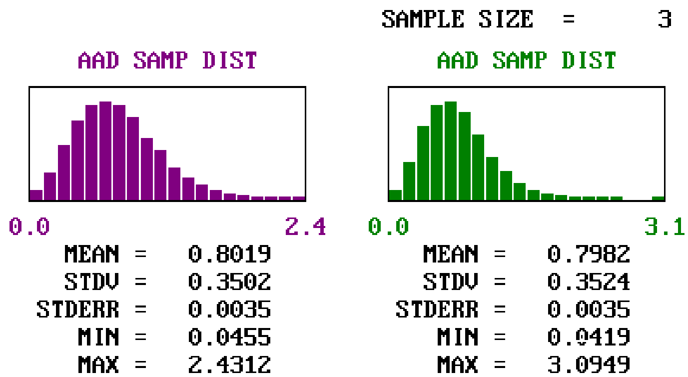


Figure 36b. Comparison of the shapes and spread of estimated AAD values for populations centered on the target and with various skewness coefficients, sample size=3.

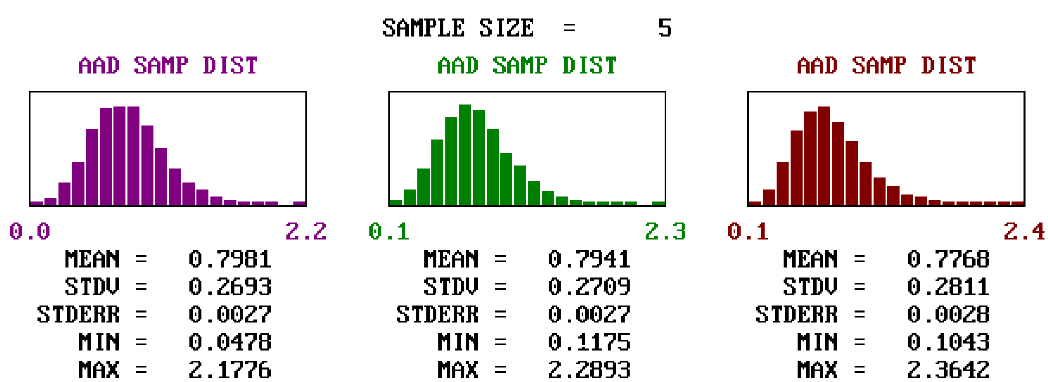


Figure 36c. Comparison of the shapes and spread of estimated AAD values for populations centered on the target and with various skewness coefficients, sample size = 5.

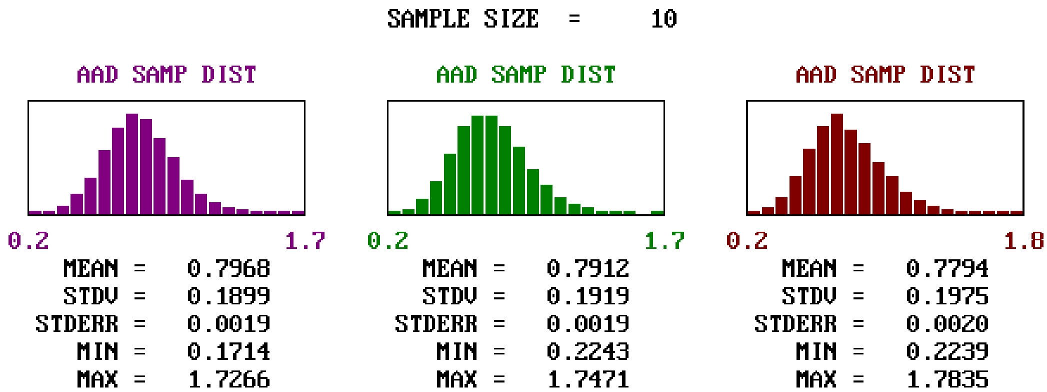


Figure 36d. Comparison of the shapes and spread of estimated AAD values for populations centered on the target and with various skewness coefficients, sample size = 10.

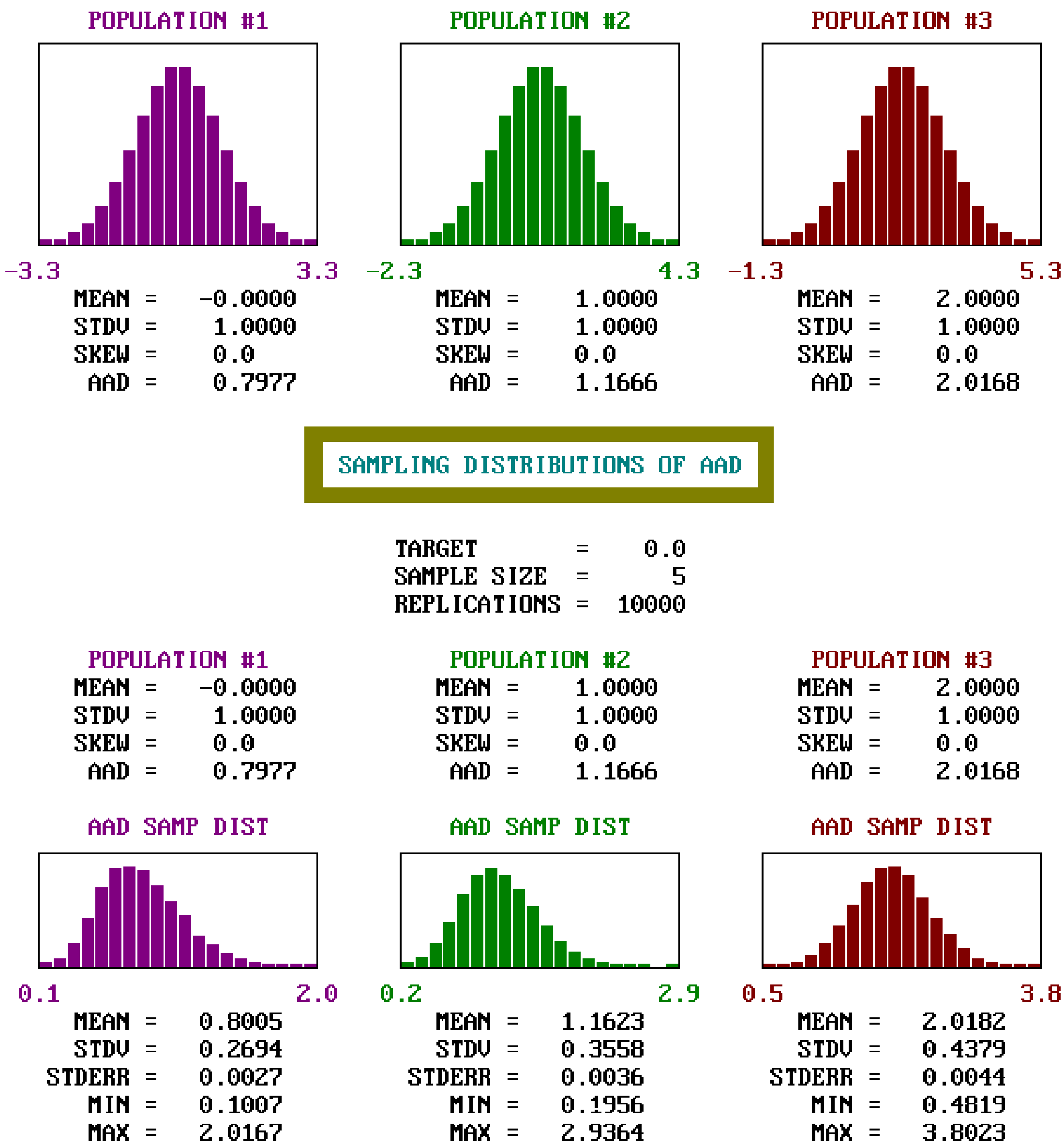


Figure 37. Comparison of the shapes and spread of estimated AAD values for normal populations centered on and offset from the target.

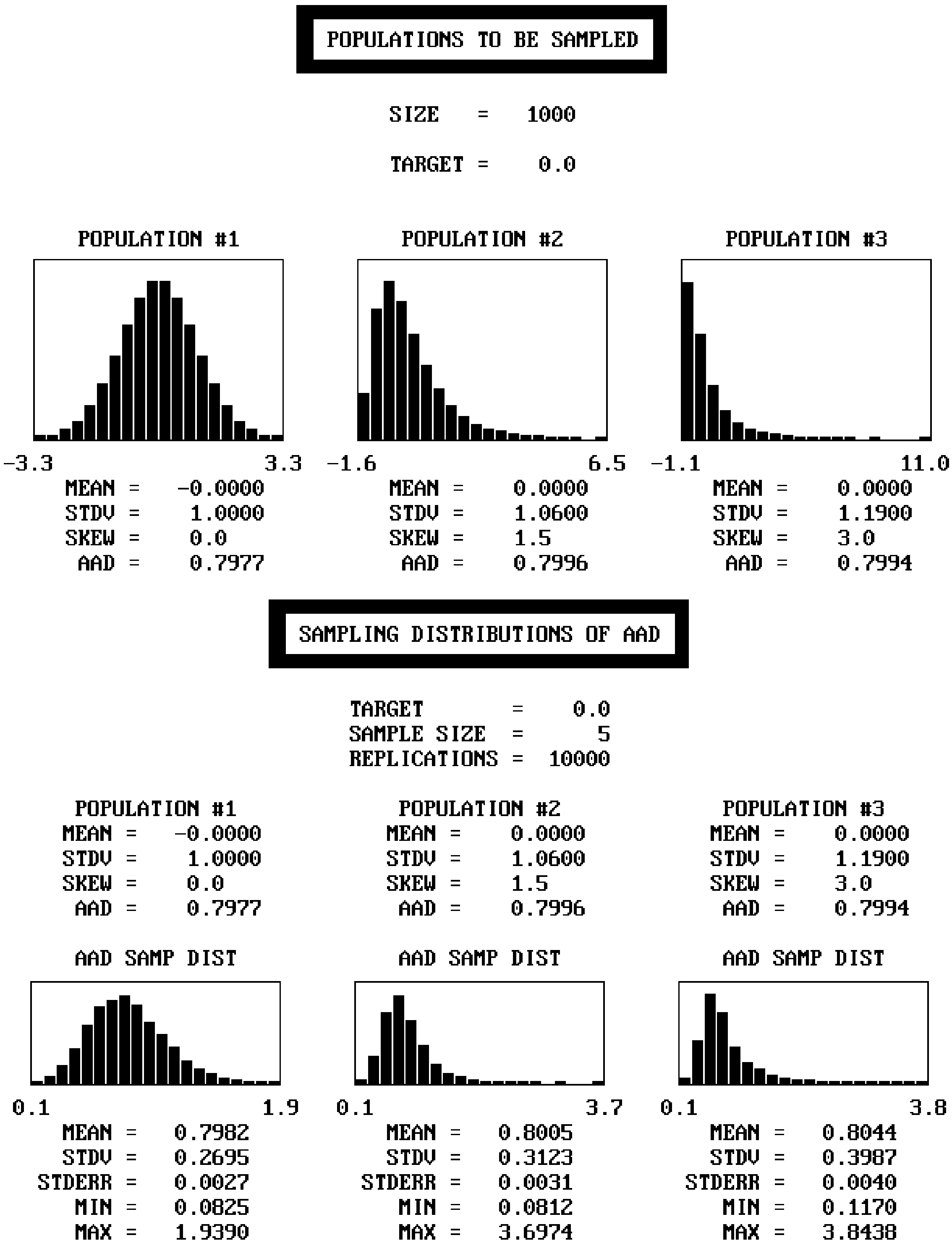


Figure 38. Example of populations that are very dissimilar in shape, but have approximately the same AAD.

There are obviously an infinite number of combinations of population mean offsets from the target and skewness coefficients. It is not possible to evaluate all of these combinations. However, it is possible to hold the others constant and change only one variable to see what effect this might have.

For sample sizes = 3, 5, and 10, table 24 lists the actual AAD values, the mean of the AAD sample values (i.e., the bias), and the standard deviation values for a number of populations that are all centered on the target, but have varying skewness levels. Figure 39 plots the results from table 24. It is apparent that the means of the AAD sample estimates are generally close to the true AAD values. In other words, the AAD estimating process appears to be unbiased. However, the standard deviations of the AAD sample estimates increase as the skewness level increases, while they decrease with increasing sample size.

For a sample size = 5, table 25 lists the actual AAD values, the mean of the AAD sample values (i.e., the bias), and the standard deviation values for a number of normal populations (i.e., skewness coefficient = 0.0) with centers that vary with respect to the amount of their offset from the target. Figure 40 plots the results from table 25. It is apparent that the means of the AAD sample estimates are generally close to the true AAD values. In other words, the AAD estimating process appears to be unbiased. However, the standard deviations of the AAD sample estimates increase as the mean offset from the target increases.

For a sample size = 5, table 26 lists the actual AAD values, the mean of the AAD sample values (i.e., the bias), and the standard deviation values for a number of normal populations that are all centered on the target, but with varying population standard deviation values. Figure 41 plots the results from table 26. It is apparent that the means of the AAD sample estimates are generally close to the true AAD values. In other words, the AAD estimating process appears to be unbiased. However, the standard deviations of the AAD sample estimates increase as the population standard deviation increases.

Table 24. Bias and spread of the AAD sample estimates for populations centered on the target, but with various levels of skewness.

n	Offset	Skewness	Std. Dev.	AAD	Bias	Std. Dev. of AAD Values
3	0.00	0.00	1.00	0.7977	0.0042	0.3502
		0.50		0.7934	0.0048	0.3524
		1.00		0.7776	0.0081	0.3659
		1.50		0.7544	0.0013	0.3783
		2.00		0.7273	–0.0067	0.3997
		2.50		0.6993	0.0061	0.4202
		3.00		0.6718	–0.0047	0.4158
5	0.00	0.00	1.00	0.7977	0.0004	0.2693
		0.50		0.7934	0.0007	0.2709
		1.00		0.7776	–0.0008	0.2811
		1.50		0.7544	–0.0026	0.2908
		2.00		0.7273	–0.0012	0.3064
		2.50		0.6993	–0.0041	0.3136
		3.00		0.6718	–0.0023	0.3316
10	0.00	0.00	1.00	0.7977	–0.0009	0.1899
		0.50		0.7934	–0.0022	0.1919
		1.00		0.7776	0.0018	0.1975
		1.50		0.7544	0.0009	0.2069
		2.00		0.7273	0.0036	0.2191
		2.50		0.6993	–0.0012	0.2233
		3.00		0.6718	0.0009	0.2334

n	=	sample size, tests per lot
Offset	=	offset of population mean from the target value, in units of standard deviation
Skewness	=	skewness coefficient for the population
Std. Dev.	=	standard deviation for the population
AAD	=	actual AAD value for the population
Bias	=	average of the estimated AAD values minus the actual AAD value
Std. Dev. of AAD Values	=	standard deviation of the estimated AAD values

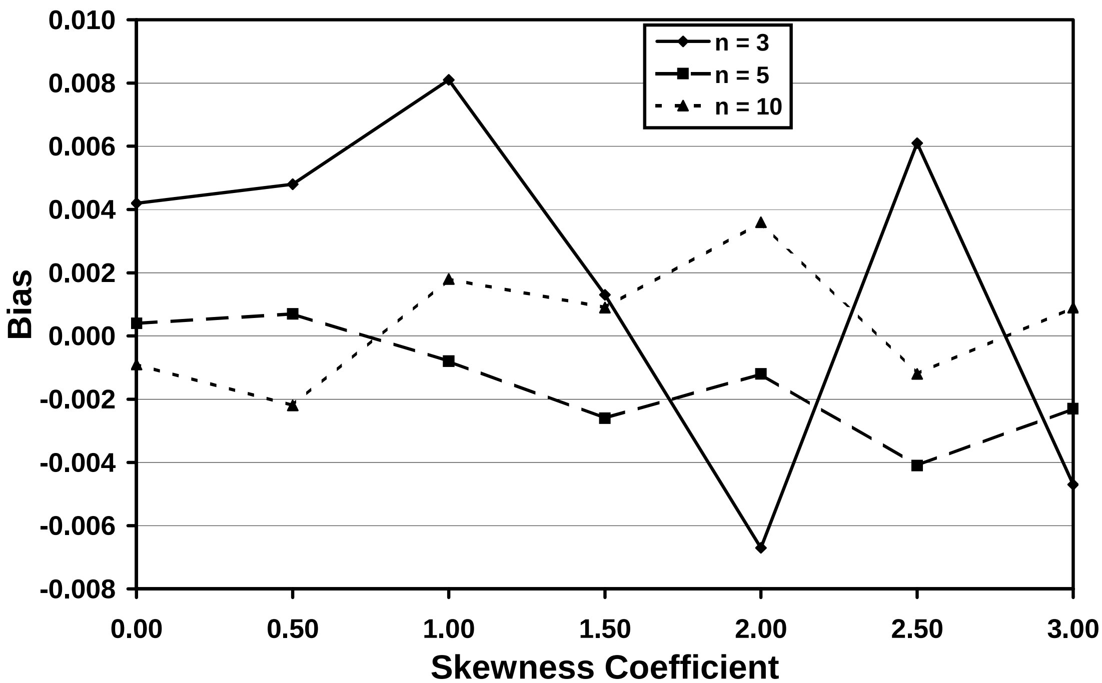


Figure 39a. Bias of the AAD sample estimates for populations centered on the target, but with various levels of skewness.

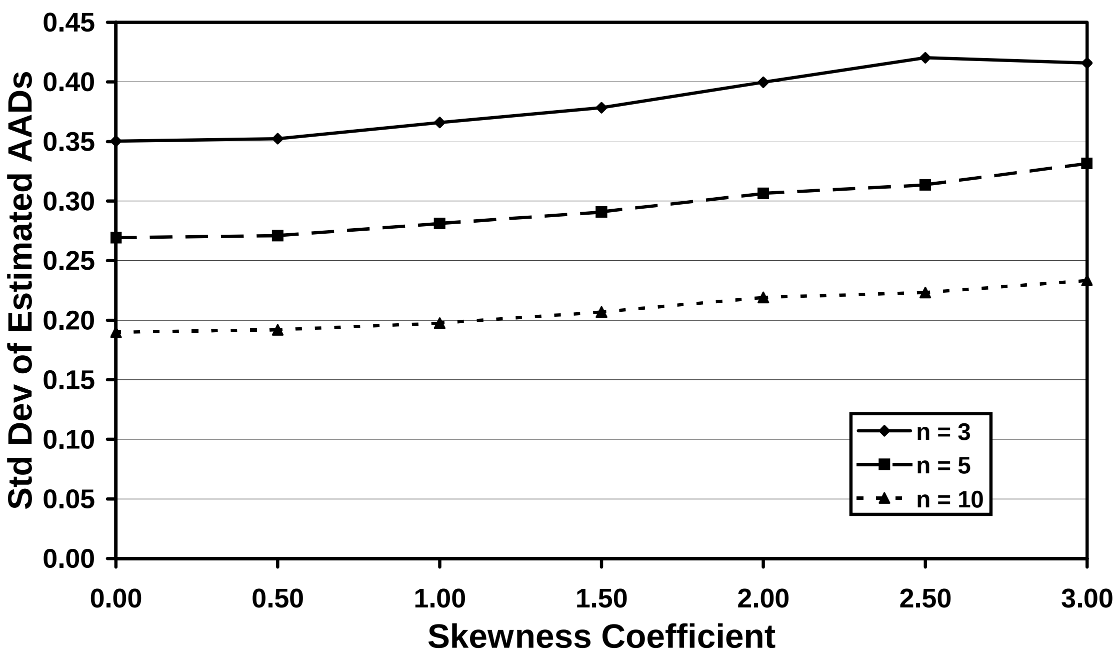


Figure 39b. Spread of the AAD sample estimates for populations centered on the target, but with various levels of skewness.

Table 25. Bias and spread of the AAD sample estimates for normal populations with various offsets from the target and n = 5.

n	Offset	Skewness	Std. Dev.	AAD	Bias	Std. Dev. of AAD Values
5	0.00	0.00	1.00	0.7977	–0.0014	0.2682
	0.25			0.8225	–0.0014	0.2808
	0.50			0.8955	0.0012	0.3000
	0.75			1.0123	0.0003	0.3258
	1.00			1.1666	0.0020	0.3590
	1.25			1.3511	–0.0030	0.3866
	1.50			1.5585	0.0037	0.4027
	1.75			1.7822	0.0007	0.4227
	2.00			2.0168	0.0024	0.4280

n	=	sample size, tests per lot
Offset	=	offset of population mean from the target value, in units of standard deviation
Skewness	=	skewness coefficient for the population
Std. Dev.	=	standard deviation for the population
AAD	=	actual AAD value for the population
Bias	=	average of the estimated AAD values minus the actual AAD value
Std. Dev. of AAD Values	=	standard deviation of the estimated AAD values

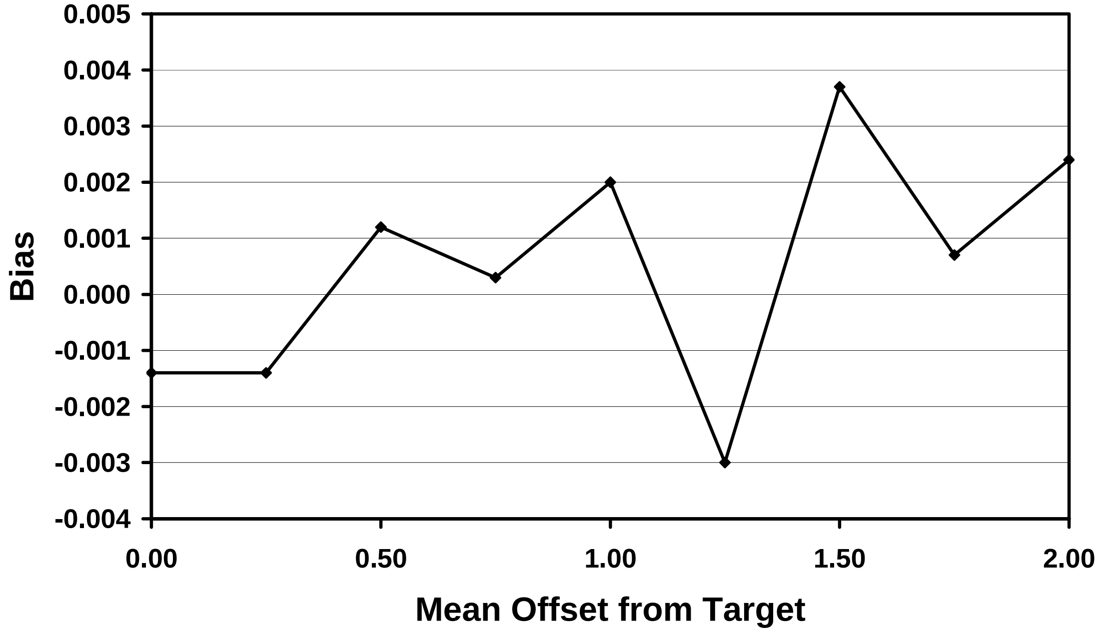


Figure 40a. Bias of the AAD samples estimates for normal populations with various offsets from the target and n=5.

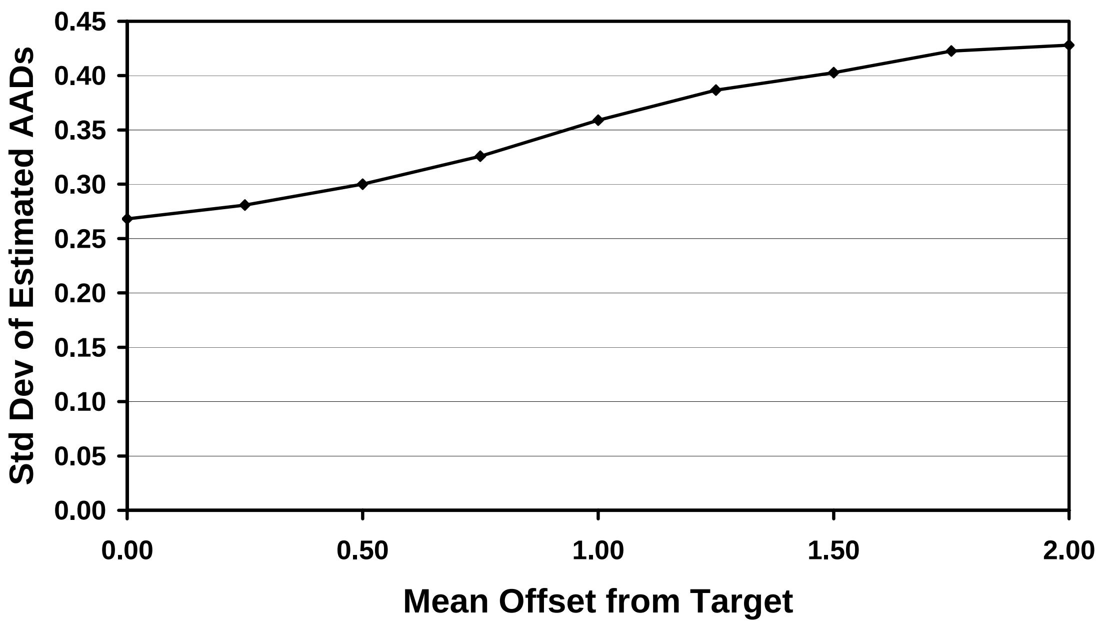


Figure 40b. Spread of the Estimated AAD sample estimates for normal populations with various offsets from the target and *n* = 5.

Table 26. Bias and spread of the AAD sample estimates for normal populations with various standard deviation values and n = 5.

n	Offset	Skewness	Std. Dev.	AAD	Bias	Std. Dev. of AAD Values
5	0.00	0.00	1.00	0.7977	0.0014	0.2683
			1.25	0.9971	0.0010	0.3328
			1.50	1.1965	–0.0011	0.4037
			1.75	1.3960	0.0015	0.4784
			2.00	1.5954	–0.0037	0.5358
			2.25	1.7948	–0.0007	0.6050
			2.50	1.9942	0.0055	0.6773
			2.75	2.1936	0.0001	0.7341
			3.00	2.3931	0.0012	0.8164

n = sample size, tests per lot
Offset = offset of population mean from the target value, in units of standard deviation
Skewness = skewness coefficient for the population
Std. Dev. = standard deviation for the population
AAD = actual AAD value for the population
Bias = average of the estimated AAD values minus the actual AAD value
Std. Dev. of AAD Values = standard deviation of the estimated AAD values

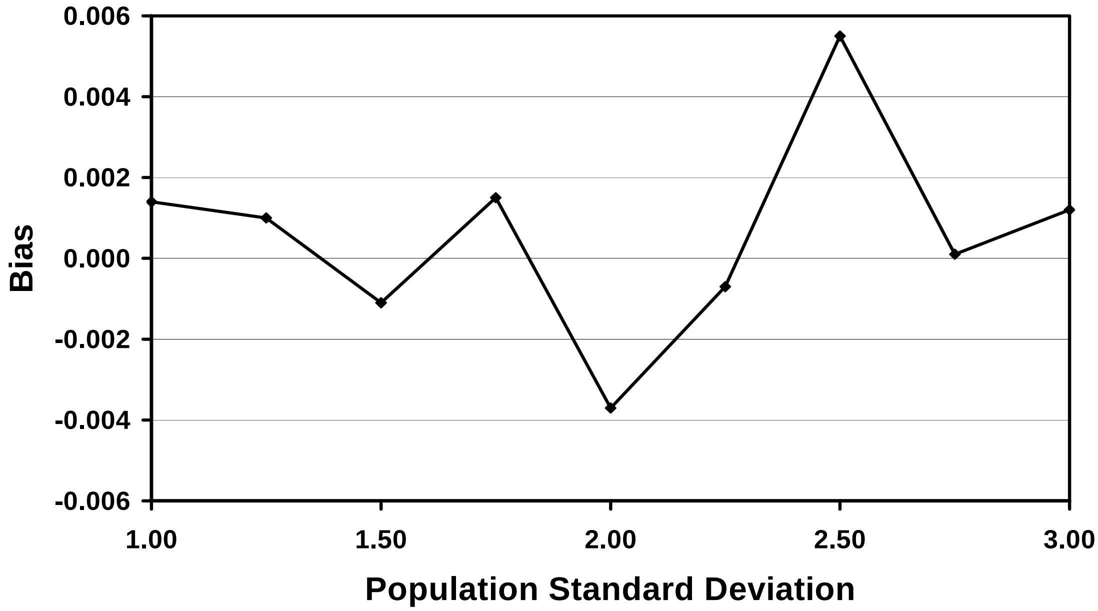


Figure 41a. Bias of the AAD sample estimates for normal populations centered on the target, with various standard deviation values, and n=5.

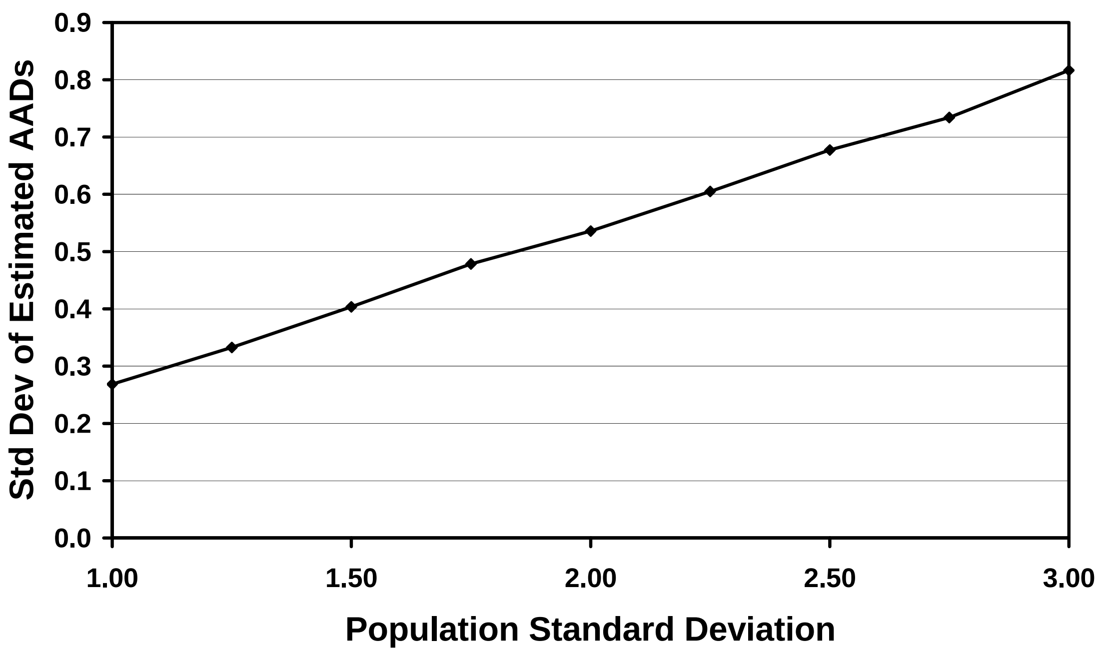


Figure 41b. Standard Deviation of the AAD sample estimates for normal populations centered on the target, with various standard deviation values, and *n* = 5.

Comparison of PWL and AAD

Both PWL and AAD seemed to perform reasonably well when used with skewed distributions. They both provided bias estimates that were close to the actual values. They both had decreasing variability of the estimated values as the sample size increased. They both had increasing variability of the estimated values as the departures from normality became more pronounced.

However, it is difficult to compare the performance of the two since the magnitude of the numbers that are calculated varies so greatly. The PWL (or PD) values can range from 0 to 100, whereas AAD values will typically range from 0 to a value of less than 10. Also, if the PWL and AAD results were to be converted to payment factors by some equation, the multiplier for the AAD equation is likely to be much larger than the one for the PWL equation.

While no definitive conclusion can be drawn, both methods will probably provide satisfactory results for the amount of skewness that is likely to be found in highway materials applications.

BIMODAL DISTRIBUTIONS

It is possible that a process that is believed to be normally distributed could change, thereby resulting in a bimodal distribution. This could happen, for instance, if the asphalt content was changed at a mixing plant and the data from before and after the change were considered in the same lot calculations. Another way in which a bimodal distribution could appear is if there were two different plants providing materials for the same project. If the plants had different process means, then the combined output from the two plants might result in a bimodal distribution.

PWL Evaluation

A computer simulation program was developed to investigate the performance of the PWL estimating procedures when working with a bimodal distribution. The program forms a bimodal distribution in the same manner in which it would happen in actual practice—by combining two different normal populations. The output screen for the program provides a plot of the distribution of each of the individual populations and a plot for the combined distribution.

When two populations are combined, they can differ in three ways:

- Different means, but the same standard deviations.
- Same means, but different standard deviations.
- Both different means and different standard deviations.

The program can handle any of these three situations. Each of these three scenarios is considered in the following sections:

Different Means: A sample output screen from the program is shown in figure 42. In this example, two identical populations are combined. Each population, shown as Distribution #1 and Distribution #2 in figure 42, has a mean of 10.00 and a standard deviation of 1.00. Each population also has a PD of 9.99. The top two plots in figure 42 represent these populations. The combined population has a mean = 10.00, standard deviation = 1.00, and PD = 9.99. Figure 42 also shows the lower specification limit, Limit1, to be 8.72 and the upper specification limit, Limit2, to be 999.00. These limits are input by the user. Inputting a very large upper limit that will never be reached makes this an example for a specification item that has only a lower limit. By trial and error, input limits can be easily developed to represent any value for the population PD (or PWL) value. The bottom plot in figure 42 represents the combined population.

The right side of the screen shows the results of the simulation of, in this example, 1000 replications of a sample size = 5. The average of these 1000 sample PD values is shown as 10.16, with a standard deviation of 11.46. A histogram of the differences between the individual sample PD values and the true population PD value is also shown. In the bottom right corner of the screen, Summary of Results shows that the PD bias was 0.16 and that we are 95-percent confident that the true value falls within the range of 0.16 ±0.71.

It is to be expected that if two identical distributions are combined, the result will be the same distribution. What will happen then if distributions with different means, but the same standard deviations, are combined? Figure 43 shows, for a sample size = 5, the output screens for cases where the offset between the population means varies from 1 to 5 standard deviations. Note that for 1 and 2 standard deviation offsets, the combined populations still appear to be approximately normal. Not until the mean offset reaches 3 does the bimodal shape begin to become apparent. It is unlikely that distributions that are so different would be combined, and if they were combined, problems would probably arise that would make the error obvious. This means that for most practical cases, if two distributions with the same standard deviation were combined, the combined distribution would not differ enough from normal to adversely affect the PD or PWL estimates.

Table 27 presents the results of simulation analyses in which two distributions with equal standard deviations and a variety of mean offsets were combined. This table represents the case of a single specification limit. Table 28 provides similar results for simulation analyses of two-sided specifications. The tables indicate that the average bias values are reasonably small, particularly for mean offsets of up to 2 or 3 standard deviations.

Different Standard Deviations: The combining of two normal distributions with different standard deviations, but with equal means, presents no different situation with respect to estimating PWL than does a single normal distribution. This is because whenever two normal distributions with equal means are combined, the resulting distribution will always be another normal distribution, regardless of the values of their standard deviations. This is illustrated in figure 44, which shows example output screens in which several pairs of normal distributions with differing standard deviations are combined. The resulting combined distribution always has the same mean as the two initial distributions and the only difference is the standard deviation for the combined distribution. Therefore, no additional discussion is necessary for the case of combining normal distributions with equal means.

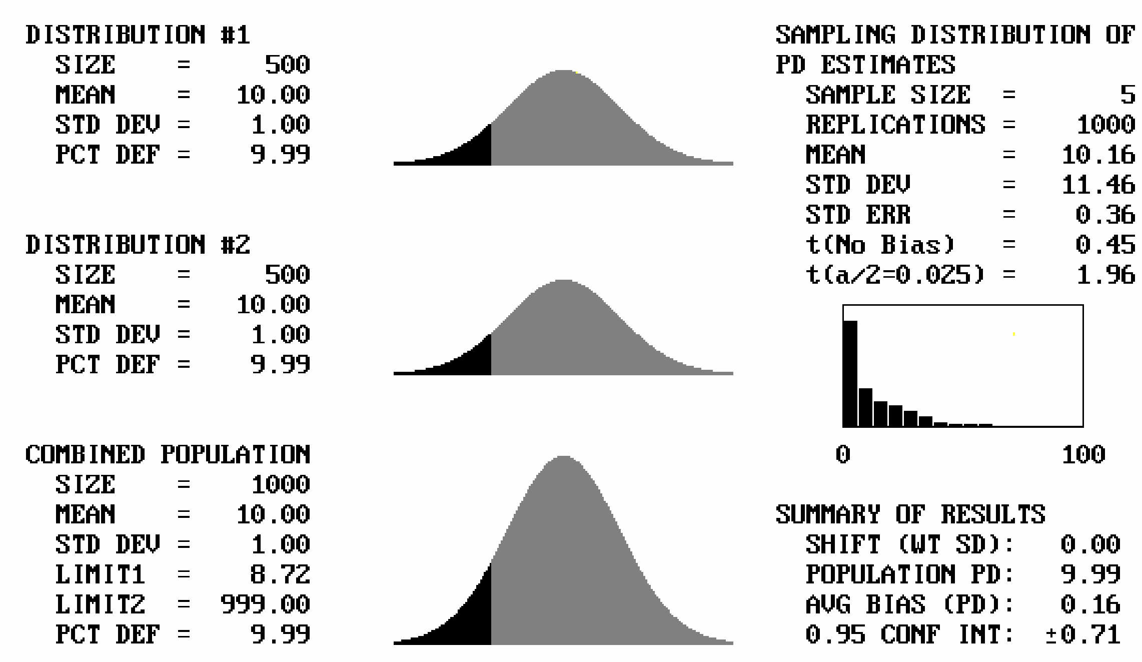


Figure 42. Sample output screen from the simulation program for bimodal distributions.

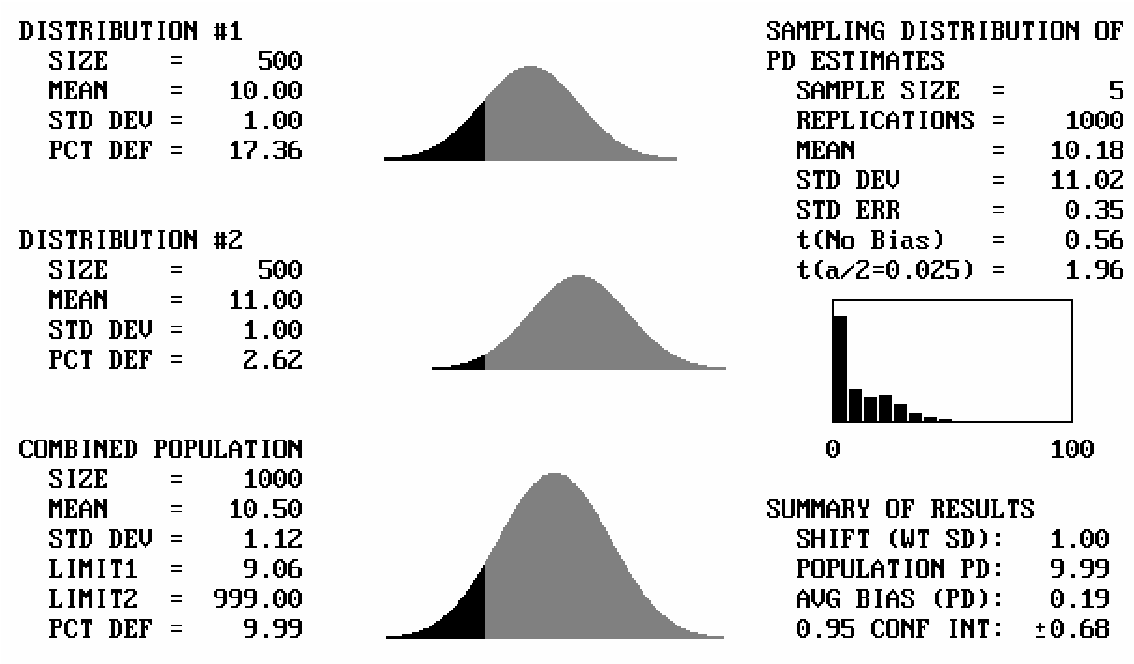


Figure 43a. Program output screens for sample size = 5 and mean offsets = 1.

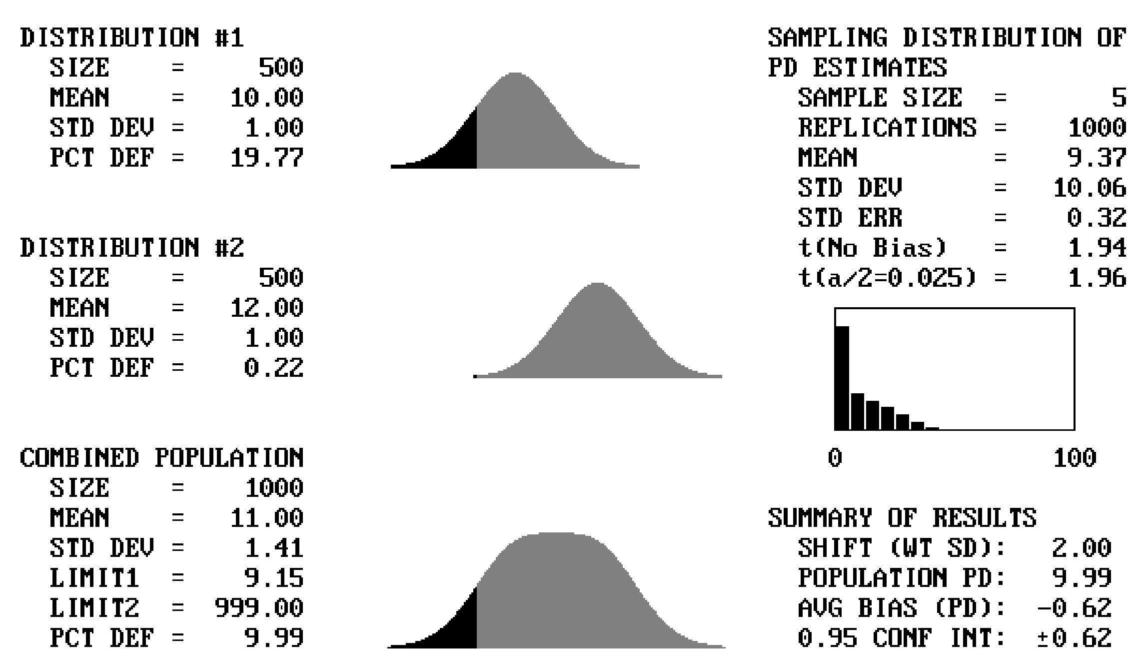


Figure 43b. Program output screens for sample size = 5 and mean offsets = 2.

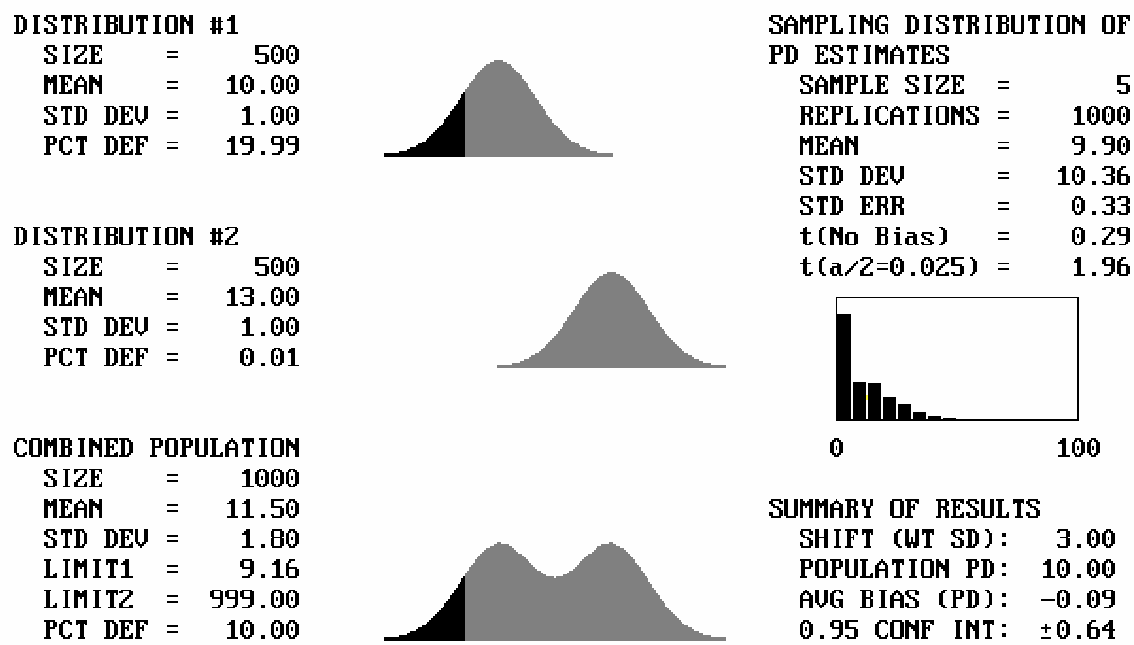


Figure 43c. Program output screens for sample size = 5 and mean offsets = 3.

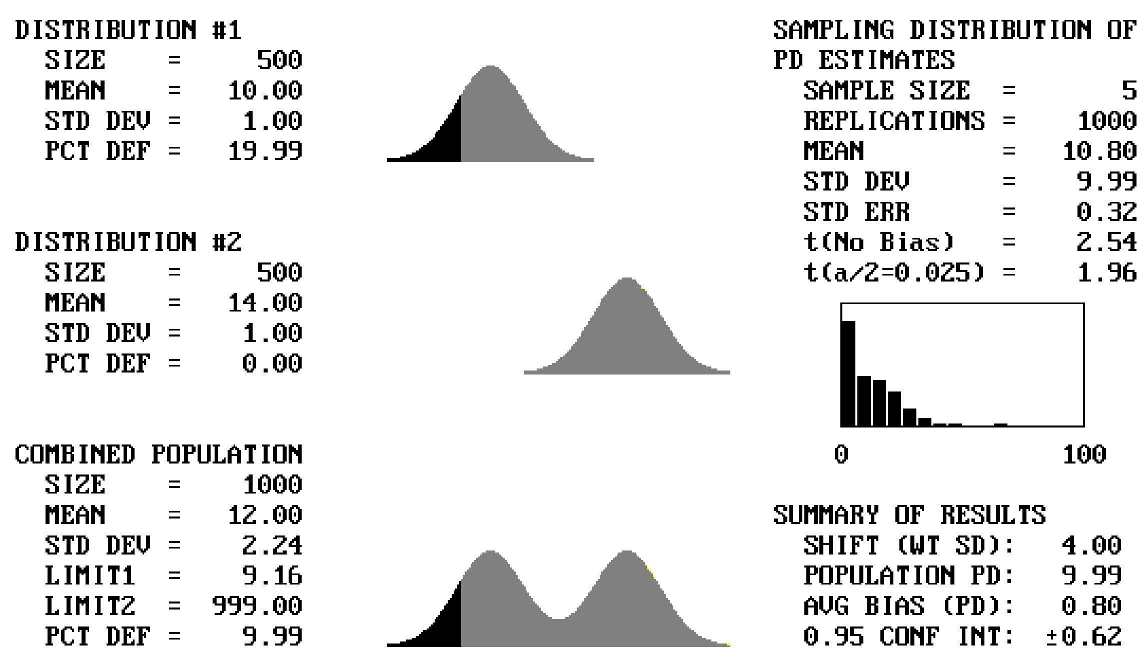


Figure 43d. Program output screens for sample size = 5 and mean offsets = 4.

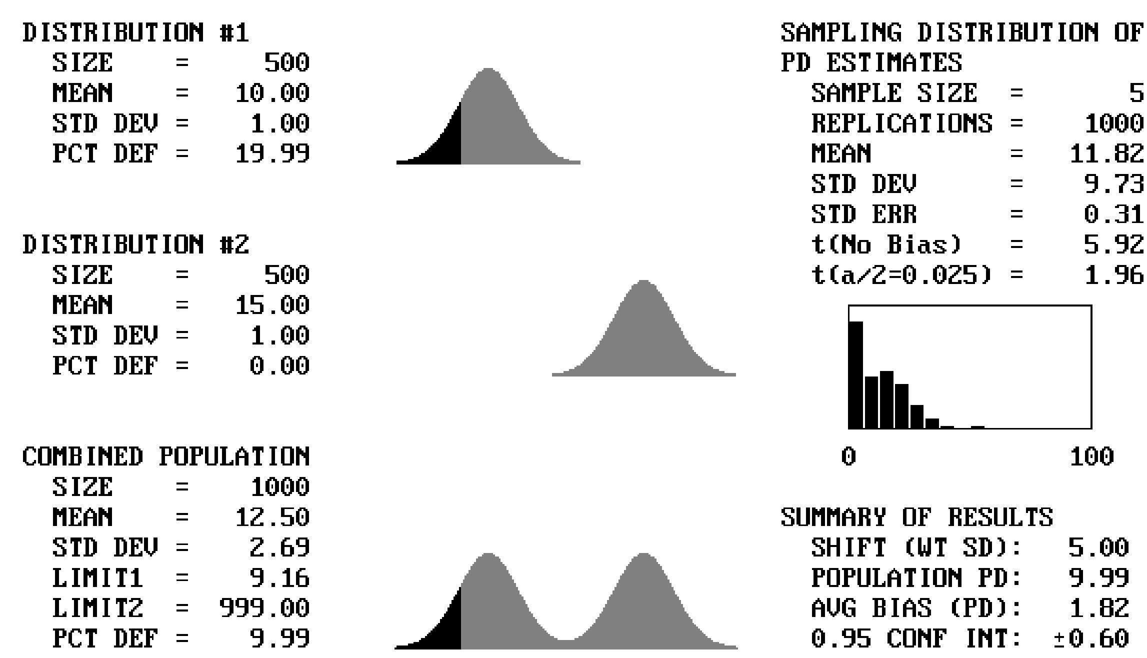


Figure 43e. Program output screens for sample size = 5 and mean offsets = 5.

Table 27. Bias results for combining two populations with equal σ with one-sided limits.

n	Offset	Actual PD	Average PD	Average Bias	95% CI	Different?
5	0.0	9.99	10.16	+0.16	±0.71	N
5	1.0	9.99	10.18	+0.19	±0.68	N
5	2.0	9.99	9.37	–0.62	±0.62	N
5	3.0	10.00	9.90	–0.09	±0.64	N
5	4.0	9.99	10.80	+0.80	±0.62	S
5	5.0	9.99	11.82	+1.82	±0.60	S
3	0.0	9.99	10.00	+0.01	±0.97	N
3	1.0	9.99	10.32	+0.33	±0.94	N
3	2.0	9.99	9.25	–0.74	±0.89	N
3	3.0	10.00	9.74	–0.26	±0.88	N
3	4.0	9.99	10.57	+0.58	±0.92	N
3	5.0	9.99	11.85	+1.86	±0.89	S
5	0.0	30.02	29.62	–0.39	±1.05	N
5	1.0	30.01	29.25	–0.75	±1.06	N
5	2.0	30.01	27.27	–1.74	±1.09	S
5	3.0	30.01	26.40	–3.61	±0.99	S
5	4.0	30.00	24.67	–5.33	±0.99	S
5	5.0	30.01	23.17	–6.84	±0.94	S
3	0.0	30.02	29.91	–0.10	±1.54	N
3	1.0	30.01	29.81	–0.20	±1.52	N
3	2.0	30.01	30.38	+0.37	±1.45	N
3	3.0	30.01	27.39	–2.62	±1.44	S
3	4.0	30.00	26.63	–3.37	±1.36	S
3	5.0	30.01	25.47	–4.54	±1.34	S

n = sample size
Offset = offset of the two population means in terms of the average σ of the populations
Actual PD = actual percent defective (PD) of the combined population
Average PD = average PD of 1000 simulated samples from the combined population
Average Bias = average simulated PD minus the actual PD
95% CI = interval within which 95 percent of the average bias values should fall
Different? = whether or not the average bias is significantly different from zero at the 0.05 level (N = not significantly different, S = significantly different)

Table 28. Bias results for combining two populations with equal σ with two-sided limits equidistant from the mean.

n	Offset	Actual PD	Average PD	Average Bias	95% CI	Different?
5	0.0	10.00	9.38	–0.61	±0.66	N
5	1.0	10.00	10.21	+0.21	±0.66	N
5	2.0	9.99	10.47	+0.48	±0.66	N
5	3.0	9.99	12.46	+2.47	±0.63	S
5	4.0	9.99	14.51	+4.52	±0.59	S
5	5.0	9.99	16.40	+6.41	±0.57	S
3	0.0	10.00	10.70	+0.71	±0.98	N
3	1.0	10.00	10.33	+0.34	±0.94	N
3	2.0	9.99	10.98	+0.99	±0.96	S
3	3.0	9.99	11.30	+1.31	±0.91	S
3	4.0	9.99	13.06	+3.07	±0.90	S
3	5.0	9.99	11.30	+1.31	±0.91	S
5	0.0	30.02	29.83	–0.19	±1.03	N
5	1.0	30.02	30.05	+0.03	±0.99	N
5	2.0	30.02	28.77	–1.24	±0.93	S
5	3.0	30.00	27.23	–2.77	±0.81	S
5	4.0	30.02	28.90	–1.12	±0.70	S
5	5.0	30.02	29.03	–0.99	±0.59	S
3	0.0	30.02	30.14	+0.12	±1.54	N
3	1.0	30.02	28.80	–1.21	±1.52	N
3	2.0	30.02	28.33	–1.69	±1.45	S
3	3.0	30.00	29.89	–0.11	±1.44	N
3	4.0	30.02	29.61	–0.42	±1.36	N
3	5.0	30.02	30.71	+0.70	±1.34	N

n = sample size
Offset = offset of the two population means in terms of the average σ of the populations
Actual PD = actual percent defective (PD) of the combined population
Average PD = average PD of 1000 simulated samples from the combined population
Average Bias = average simulated PD minus the actual PD
95% CI = interval within which 95 percent of the average bias values should fall
Different? = whether or not the average bias is significantly different from zero at the 0.05 level (N = not significantly different, S = significantly different)

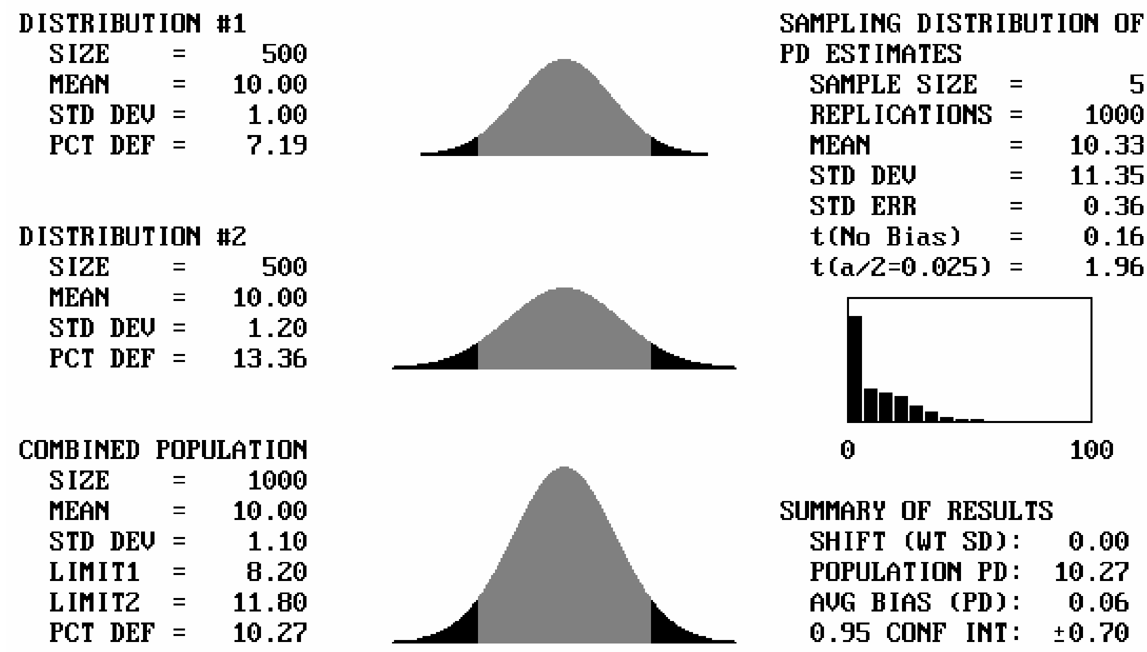


Figure 44a. Illustration 1 of program output screens when combining distributions with equal means for sample size = 5.

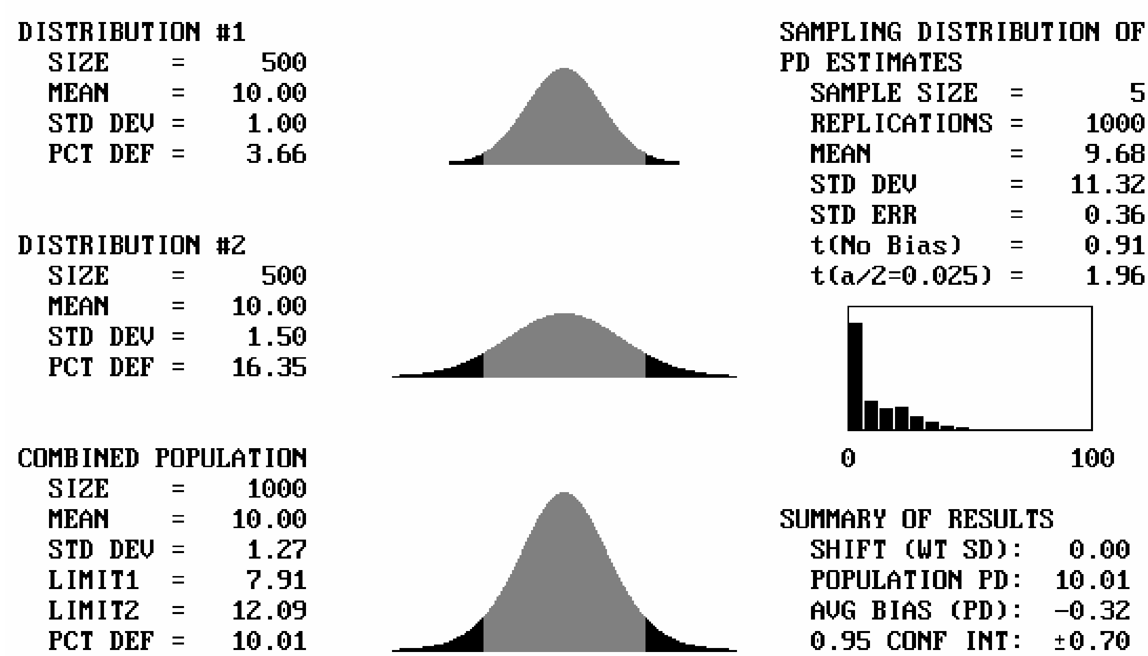


Figure 44b. Illustration 2 of program output screens when combining distributions with equal means for sample size = 5.

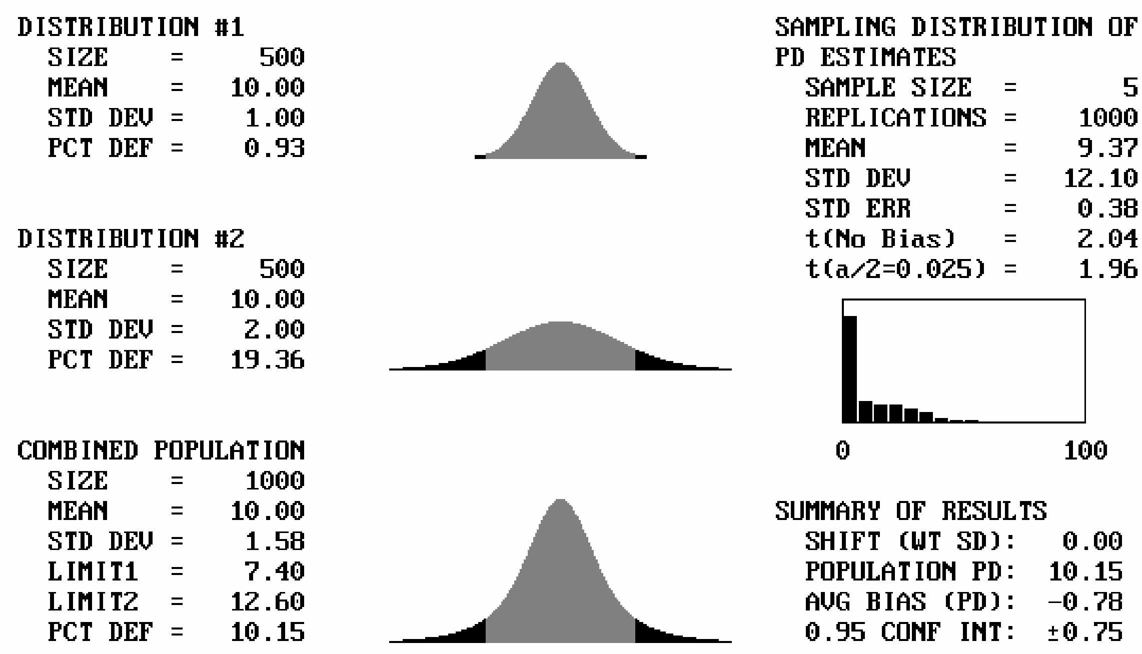


Figure 44c. Illustration 3 of program output screens when combining distributions with equal means for sample size = 5.

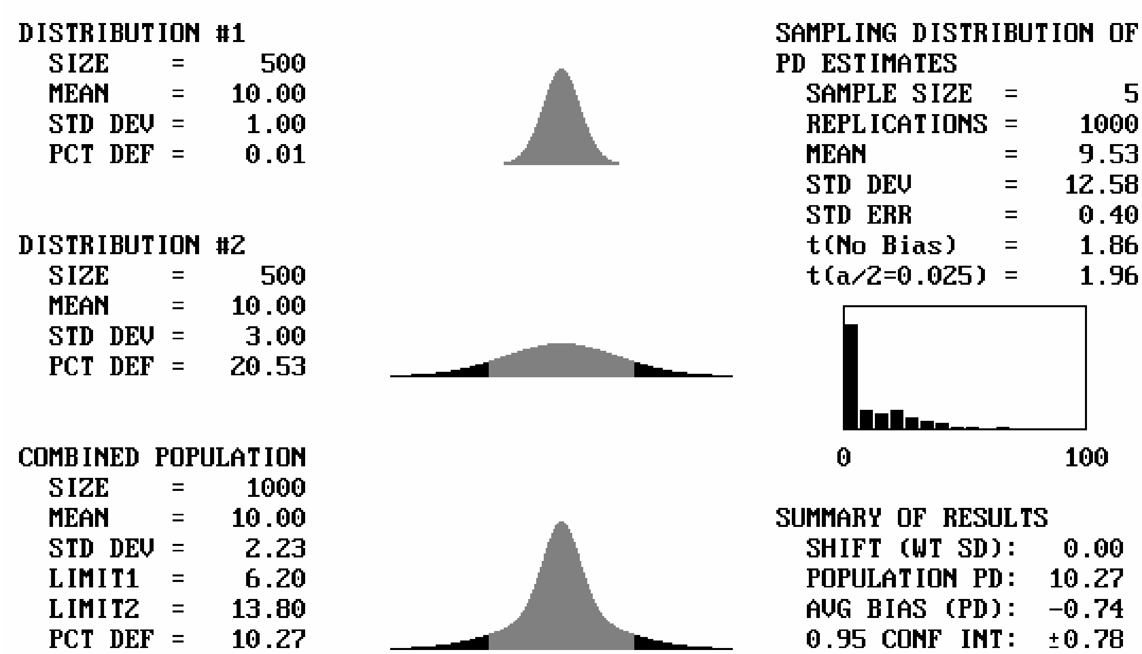


Figure 44d. Illustration 4 of program output screens when combining distributions with equal means for sample size = 5.

Different Means and Different Standard Deviations: When populations with different means and different standard deviations are combined, the shape of the combined distribution depends on the amount of difference between them. This is illustrated in the combined distribution shapes in table 29. The original distributions in this figure are different in both mean and standard deviation. The ratio of the standard deviations of the two populations is 2.0. The offset of the two means is stated in terms of the *weighted* standard deviation. This is simply the arithmetic average of the standard deviations for the two distributions that are combined.

For small offsets, the combined distribution is simply a unimodal skewed distribution. It is not until the weighted offset reaches 2.0 that a slight bulge appears in the long tail of the skewed distribution. This bulge becomes more distinct and the combined distribution approaches a true bimodal shape as the weighted offset increases to above 3.0.

To investigate the performance of the PWL estimator when two different distributions are combined, the simulation program was used to generate 10,000 samples of size $n = 5$ from bimodal distributions with 10 PD (or 90 PWL). The PD estimate for each sample was then calculated and compared with the known actual PD of 10 (PWL of 90). The bias was then calculated as the difference between the sample PD and the actual PD. The average bias and variability for the 10,000 samples were then determined.

In the analyses, three different cases were considered:

- One-sided lower specification limit.
- Two-sided specification limit with the limits equidistant about the population mean.
- Two-sided specification limit with 5 percent of the population outside of each limit.

For each of the three specification limit cases, three scenarios were considered for the bimodal distributions. When combining the distributions, three different standard deviation relationships were used. First, equal standard deviations were used for the two distributions. This generated bimodal distributions that were symmetrical. These were used as controls against which to compare the other two cases. Two different standard deviation relationships were used to generate skewed bimodal distributions. In these two cases, the ratios of the standard deviations for the two combined populations were 1.5 and 2.0.

The *t*-statistic for a significance level of 0.05 and the 95-percent confidence interval for the average PD were calculated and used to determine whether the average bias was statistically significantly different from zero. The results of these analyses are presented in table 30. The bias in the PD (and, therefore, the PWL) estimates is more pronounced as the weighted offset between the two combined populations increases.

However, for the smaller weighted mean offsets, up to about 1.0 to 2.0, depending on the standard deviation ratio, the bias in estimating PD is reasonably small and sometimes not significantly different from zero.

Table 29. Shapes of combined distributions when the combined distributions have different means and standard deviations.

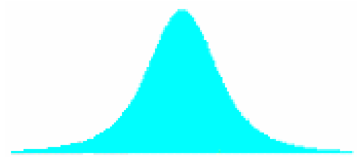
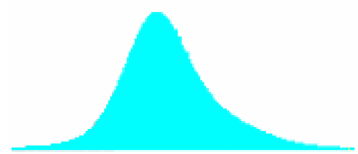
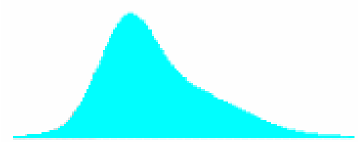
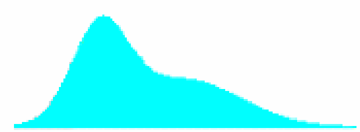
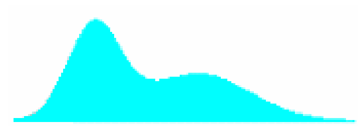
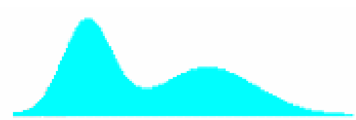
Ratio of Standard Deviations	Mean Offset, in Weighted Standard Deviation Units	Shape of Combined Distribution
2.0	0.00	
2.0	0.67	
2.0	1.33	
2.0	2.00	
2.0	2.67	
2.0	3.33	

Table 30. Bias in estimating PD when two populations are combined (sample size = 5, 10,000 simulated lots).

Std. Dev. Ratio	Weighted Mean Offset	Actual PD	Bias*	95% CI
One-Sided Specification Limits				
1.0	0.00	9.99	**0.11**	0.22
	1.00	9.99	**–0.07**	0.22
	2.00	9.99	**–0.07**	0.21
	3.00	10.00	0.24	0.20
	4.00	9.99	1.02	0.19
	5.00	9.99	1.80	0.19
1.5	0.00	10.00	**0.13**	0.23
	0.80	10.00	0.56	0.22
	1.60	10.00	1.08	0.20
	2.40	10.00	1.37	0.20
	3.20	10.00	1.93	0.19
	4.00	10.00	2.53	0.02
2.0	0.00	10.00	0.98	0.24
	0.67	10.00	1.50	0.23
	1.33	9.99	2.14	0.21
	2.00	10.00	2.37	0.20
	2.67	9.99	2.72	0.19
	3.33	10.00	3.13	0.19
Two-Sided Specification Limits—Equidistant From the Mean				
1.0	0.00	10.00	**0.05**	0.22
	1.00	10.00	**–0.02**	0.21
	2.00	9.99	0.56	0.20
	3.00	9.99	2.27	0.20
	4.00	9.99	4.61	0.19
	5.00	9.99	6.95	0.18
1.5	0.00	9.99	**–0.11**	0.22
	0.80	10.00	**0.07**	0.22
	1.60	10.00	0.57	0.22
	2.40	10.00	1.42	0.21
	3.20	10.00	2.94	0.20
	4.00	10.00	4.70	0.19
2.0	0.00	10.00	–0.37	0.24
	0.67	10.00	–0.25	0.24
	1.33	10.00	0.54	0.23
	2.00	10.00	1.16	0.23
	2.67	10.00	2.26	0.22
	3.33	10.00	3.44	0.22

*Bias values that are NOT significantly different from zero are **bolded**.

Table 30. Bias in estimating PD when two populations are combined (sample size = 5, 10,000 simulated lots) (continued).

Std. Dev. Ratio	Weighted Mean Offset	Actual PD	Bias*	95% CI
Two-Sided Specification Limits—5 PD in Each Tail				
1.0	0.00	10.00	**0.05**	0.22
	1.00	10.00	**–0.02**	0.21
	2.00	9.99	0.56	0.20
	3.00	9.99	2.27	0.20
	4.00	9.99	4.61	0.19
	5.00	9.99	6.95	0.18
1.5	0.00	9.99	**0.16**	0.23
	0.80	10.00	0.30	0.22
	1.60	10.00	0.75	0.21
	2.40	10.00	1.79	0.20
	3.20	10.00	3.23	0.20
	4.00	10.01	5.04	0.19
2.0	0.00	10.00	–0.32	0.24
	0.67	10.00	0.41	0.23
	1.33	10.00	1.14	0.22
	2.00	10.00	2.16	0.21
	2.67	10.00	2.97	0.20
	3.33	10.01	4.48	0.20

*Bias values that are NOT significantly different from zero are **bolded**.

AAD Evaluation

No AAD evaluation was conducted for bimodal distributions. The results from the PD/PWL evaluation presented above indicated that for the types of mean and standard deviation differences between populations that are likely to be combined in typical materials and construction operations, the result would be either approximately normal or unimodal skewed distributions. Since the previous AAD analyses presented above in this chapter had already evaluated the AAD bias with respect to normal and skewed distributions, it was decided that no additional AAD evaluation was necessary.

CONCLUSION

Both the PWL (and PD) estimators and the AAD estimators provided estimates with relatively little bias for the departures from normality that are likely to be encountered. Neither was obviously better or worse than the other. Based on the non-normality analysis, there is no reason to eliminate either of these quality measures, nor is there a compelling reason to select one over the other. It was decided to continue investigating both PWL and AAD as possible quality measures for use when developing payment relationships.

7. VERIFICATION PROCEDURES

INTRODUCTION

As part of the acceptance procedures and requirements, one question that must be answered is "Who is going to perform the acceptance tests?" The agency may either decide to do the acceptance testing, assign the testing to the contractor, have a combination of agency and contractor acceptance testing, or require a third party to do the testing.

The decision as to who does the testing usually emanates from the agency's personnel assessment, particularly in the days of agency downsizing. Many agencies are requiring the contractor to do the acceptance testing. This is at least partially because of agency staff reductions. What has often evolved is that the contractor is required to perform both QC and acceptance testing. If the contractor is assigned the acceptance function, the contractor's acceptance tests must be verified by the agency. The agency's verification sampling and testing function has the same underlying function as the agency's acceptance sampling and testing—to verify the quality of the product. Statistically sound verification procedures must be developed that require a separate verification program. There are several forms of verification procedures and some forms are more efficient than others. To avoid conflict, it is in the best interests of both parties to make the verification process as effective and efficient as possible.

The sources of variability are important when deciding what type of verification procedures to use. This decision depends on what the agency wants to verify. Independent samples (i.e., those obtained without respect to each other) contain up to four sources of variability: material, process, sampling, and testing. Split samples contain variability only in the testing method. Thus, if the agency wishes to verify only that the contractor's testing methods are correct, then the use of split samples is best. This is referred to as *test method verification*. If the agency wishes to verify the contractor's overall production, sampling, and testing processes, then the use of independent samples is required. This is referred to as *process verification*. Each of these types of verification is evaluated in the following sections.

HYPOTHESIS TESTING AND LEVELS OF SIGNIFICANCE

Before discussing the various procedures that can be used for test method verification or process verification, two concepts must be understood: *hypothesis testing* and *level of significance*. When it is necessary to test whether or not it is reasonable to accept an assumption about a set of data, statistical tests (called *hypothesis tests*) are conducted. Strictly speaking, a statistical test neither proves nor disproves a hypothesis. What it does is prescribe a formal manner in which evidence is to be examined to make a decision regarding whether or not the hypothesis is correct.

To perform a hypothesis test, it is first necessary to define an assumed set of conditions known as the *null hypothesis* (H_0). Additionally, an alternative hypothesis (H_a) is, as the name implies, an alternative set of conditions that will be assumed to exist if the null hypothesis is rejected. The statistical procedure consists of assuming that the null hypothesis is true and then examining the data to see if there is sufficient evidence that it should be rejected. The H_0 cannot actually be proved, only disproved. If the null hypothesis cannot be disproved (or, to be statistically correct, rejected), it should be stated that we *fail to reject*, rather than *prove* or *accept*, the hypothesis. In practice, some people use *accept* rather than *fail to reject*, although this is not exactly statistically correct.

Verification testing is simply hypothesis testing. For test method or process verification purposes, the null hypothesis would be that the contractor's tests and the agency's tests have equal means, while the alternate hypothesis would be that the means are not equal.

Hypothesis tests are conducted at a selected level of significance, α, where α is the probability of incorrectly rejecting the H_0 when it is actually true. The value of α is typically selected as 0.10, 0.05, or 0.01. For example, if $\alpha = 0.01$ and the null hypothesis is rejected, then there is only 1 chance in 100 that H_0 is true and was rejected in error.

The performance of hypothesis tests, or verification tests, can be evaluated by using OC curves. OC curves plot either the probability of not detecting a difference (i.e., accepting the null hypothesis that the populations are equal) or the probability of detecting a difference (i.e., rejecting the null hypothesis that the populations are equal) versus the actual difference between the two populations being compared. Curves that plot the probability of detecting a difference are sometimes call *power curves* because they plot the power of the statistical test procedure to detect a given difference.

Just as there is a risk of incorrectly rejecting the H_0 when it is actually true, which is called the *type I* (or α) *error*, there is also a risk of failing to reject the H_0 when it is actually false. This is called the *type II* (or β) *error*. The *power* is the probability of rejecting the H_0 when it is actually false and it is equal to $1 - \beta$. Both α and β are important and are used with the OC curves when determining the appropriate sample size to be used.

TEST METHOD VERIFICATION

The procedures for verifying the testing procedures should be based on split samples so that the testing method is the only source of variability present. The two procedures used most often for test method verification are: (1) comparing the difference between the split-sample results to a maximum allowable difference, and (2) the use of the *t*-test for paired measurements (i.e., the paired *t*-test). In this report, these are referred to as the *maximum allowable difference* and the *paired t-test*, respectively, and each is discussed below.

Maximum Allowable Difference

This is the simplest procedure that can be used for verification, although it is the least powerful. In this method, usually a single sample is split into two portions, with one portion tested by the contractor and the other portion tested by the agency. The difference between the two test results is then compared to a maximum allowable difference. Because the procedure uses only two test results, it cannot detect real differences unless the results are far apart.

The value selected for the maximum allowable difference is usually selected in the same manner as the D2S limits contained in many American Association of State Highway and Transportation Officials (AASHTO) and American Society for Testing and Materials (ASTM) test procedures. The D2S limit indicates the maximum acceptable difference between two results obtained on test portions of the same material (and thus applies only to split samples) and is provided for single- and multi-laboratory situations. It represents the difference between two individual test results that has approximately a 5-percent chance of being exceeded if the tests are actually from the same population.

Stated in general statistical terminology, the maximum allowable difference is set at two times the standard deviation of the distribution of the differences that would be obtained if the two test populations (the contractor's and the agency's) were actually equal. In other words, if the two populations are truly the same, there is approximately a 0.05 chance that this verification method will find them to be not equal. Therefore, the level of significance is 0.05 (5 percent).

OC Curves: OC curves were developed to evaluate the performance of the maximum allowable difference method for test method verification. In this method, a test is performed on a single split sample to compare the agency's and the contractor's test results. If we assume that both of these split test results are from normally distributed subpopulations, then we can calculate the variance of the difference and use it to calculate two standard deviation limits (approximately 95 percent) for the sample difference quantity.

Suppose that the agency's subpopulation has a variance σ_A^2 and the contractor's subpopulation has a variance σ_C^2. Since the variance of the difference in two independent random variables is the sum of the variances, the variance of the difference in an agency's observation and a contractor's observation is $\sigma_A^2 + \sigma_C^2$. The maximum allowable difference is based on the test standard deviation, which may be provided in the form of D2S limits. Let us call this test standard deviation σ_{test}. Under an assumption that $\sigma_A^2 = \sigma_C^2 = \sigma_{test}^2$, this variance of a difference becomes $2\sigma_{test}^2$.

The maximum allowable difference limits are set as two times the standard deviation of the test differences (i.e., approximately 95-percent limits). This, therefore, sets the limits at $\pm 2\sqrt{2\sigma_{test}^2}$, which is $\pm 2\sqrt{2}\sigma_{test}$ (or $\pm 2.8284\,\sigma_{test}$). Without loss of generality, we can assume $\sigma_{test} = 1$, along with an assumption of a mean difference of 0, and use the standard normal distribution with a region between –2.8284 and +2.8284 as the acceptance region for the difference in an agency's test result and a contractor's test result. With these two limits fixed, we can calculate the power of this decisionmaking process relative to various true differences in the underlying subpopulation means and/or various ratios of the true underlying subpopulation standard deviations.

These power values can conveniently be displayed as a three-dimensional surface. If we vary the mean difference along the first axis and the standard deviation ratio along a second axis, we can show power on the vertical axis. The agency's subpopulation, the contractor's subpopulation, or both, could have standard deviations that are smaller, about the same, or larger than the supplied

σ_{test} value. To develop OC curves, these situations were represented in terms of the minimum standard deviation between the contractor's population and the agency's population as follows:

- Minimum standard deviation equals the test standard deviation (σ_{test}).
- Minimum standard deviation equals half the test standard deviation.
- Minimum standard deviation equals twice the test standard deviation.

Figures 45 through 47 show the OC curves for each of the above cases. The power values are shown where the ratio of the larger of the agency's or the contractor's standard deviation to the smaller of the agency's or contractor's standard deviation is varied over the values 0, 1, 2, 3, 4, and 5. The mean difference given along the horizontal axis (values of 0, 1, 2, and 3) represents the difference in the agency's and contractor's subpopulation means expressed as multiples of σ_{test}.

In figure 45, which shows the case when the minimum standard deviation equals the test standard deviation (σ_{test}), even when the ratio of the contractor's and agency's standard deviations is 5 and the difference between the contractor's and the agency's means is three times the value for σ_{test}, there is less than a 70-percent chance of detecting the difference based on the results from a single split sample. As would be expected, the power values decrease when the minimum standard deviation is half of σ_{test} (figure 46) and increase when the minimum standard deviation is twice σ_{test} (figure 47).

As is the case with any method based on a sample size = 1, the D2S method does not have much power to detect the differences between the contractor's and the agency's populations. The appeal of the maximum allowable difference method lies in its simplicity, rather than in its power.

Average Run Length: The maximum allowable difference method was also evaluated based on the average run length. The average run length is the average number of lots that it takes to identify a difference between dissimilar populations. As such, the shorter the average run length, the better.

Various actual differences between the contractor's and the agency's population means and standard deviations were considered in the analysis. In the results that are presented, *i* refers to the difference (in units of the agency's population standard deviation) between the agency's and the contractor's population means. Also, *j* refers to the ratio of the contractor's population standard deviation to the agency's population standard deviation. In the analyses, *i* values of 0, 1, 2, and 3 were used, while the *j* values used were 0.5, 1.0, 1.5, and 2.0. Some examples of these *i* and *j* values are illustrated in figure 48.

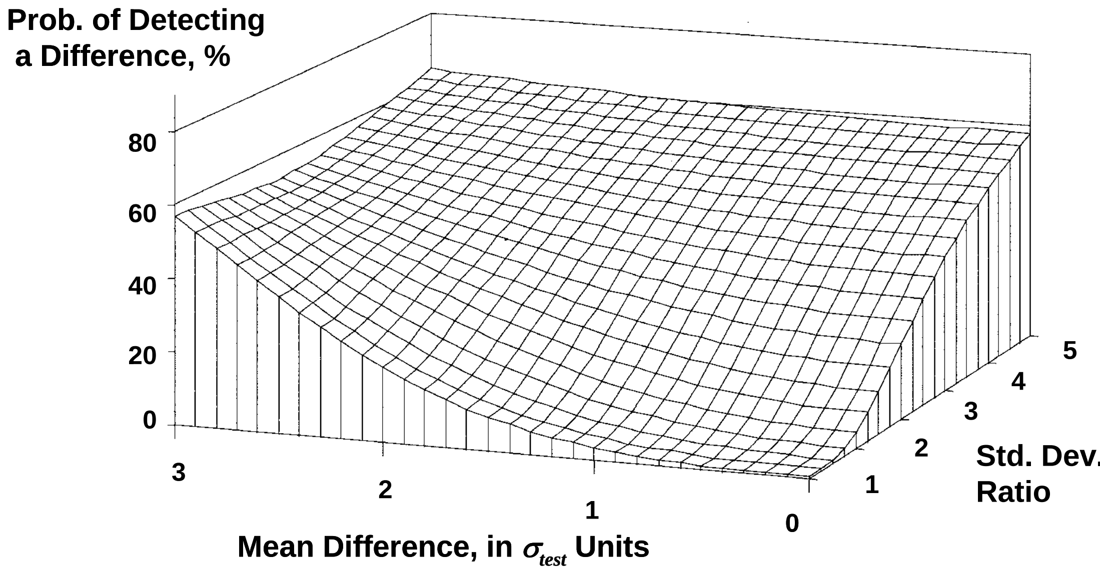


Figure 45. OC surface for the maximum allowable difference test method verification method (assuming the smaller $\sigma = \sigma_{test}$).

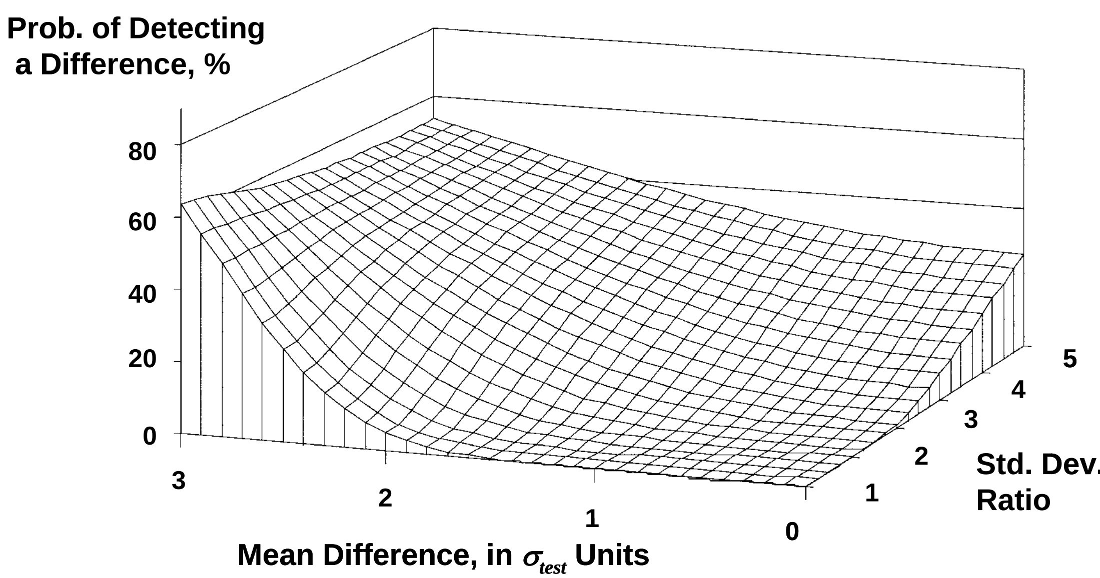


Figure 46. OC surface for the maximum allowable difference test method verification method (assuming the smaller $\sigma = 0.5\sigma_{test}$).

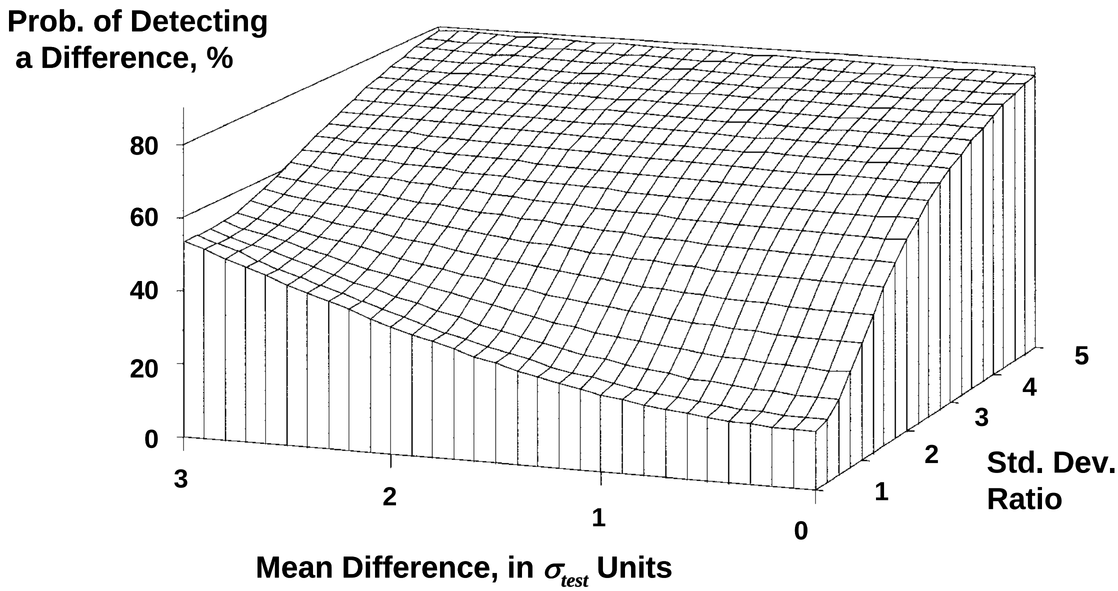


Figure 47. OC surface for the maximum allowable difference test method verification method (assuming the smaller $\sigma = 2\sigma_{test}$).

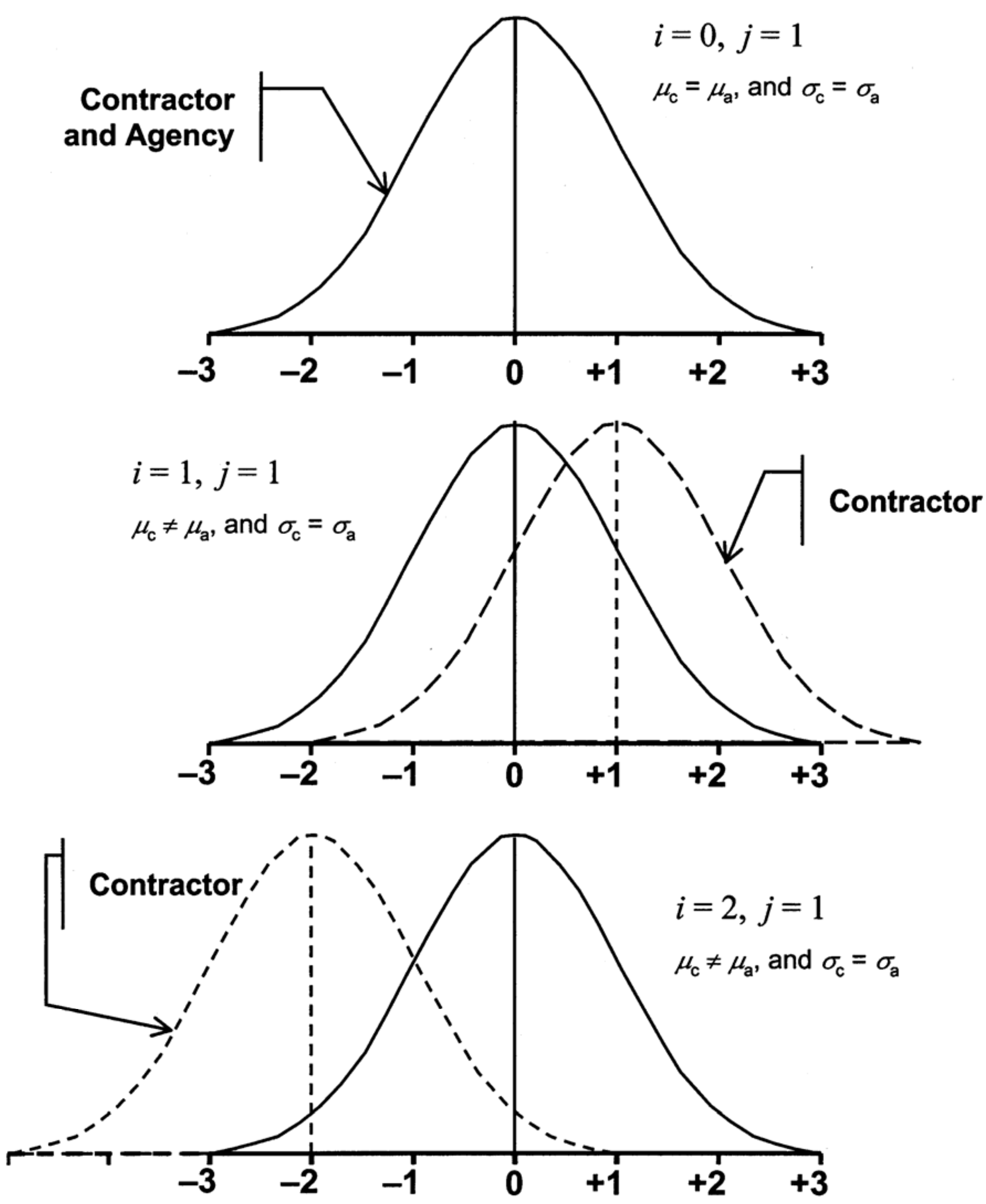
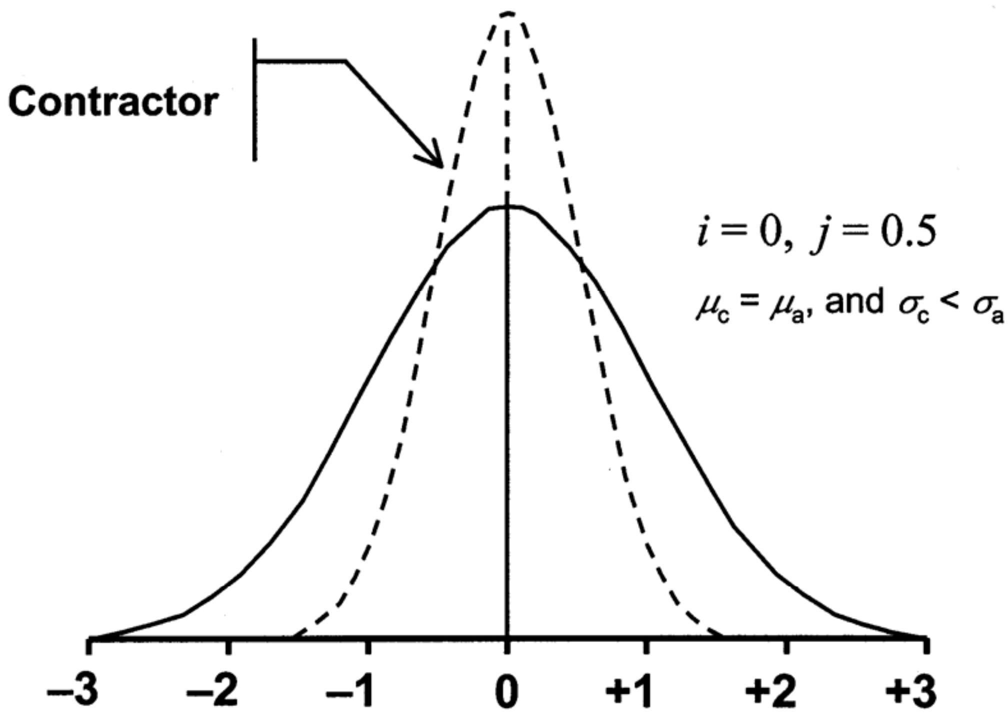


Figure 48a. Example 1 of some of the cases considered in the average run length analysis for the maximum allowable difference method.



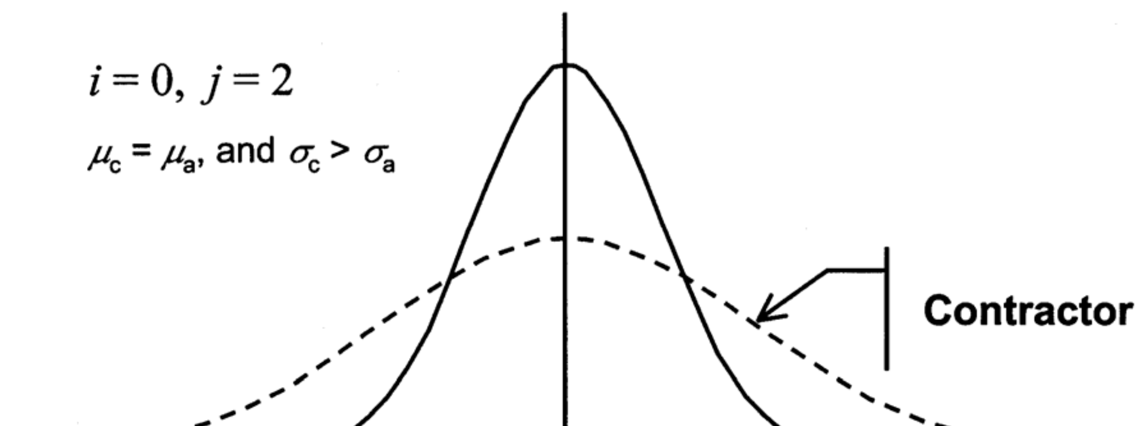


Figure 48b. Example 2 of some of the cases considered in the average run length analysis for the maximum allowable difference method.

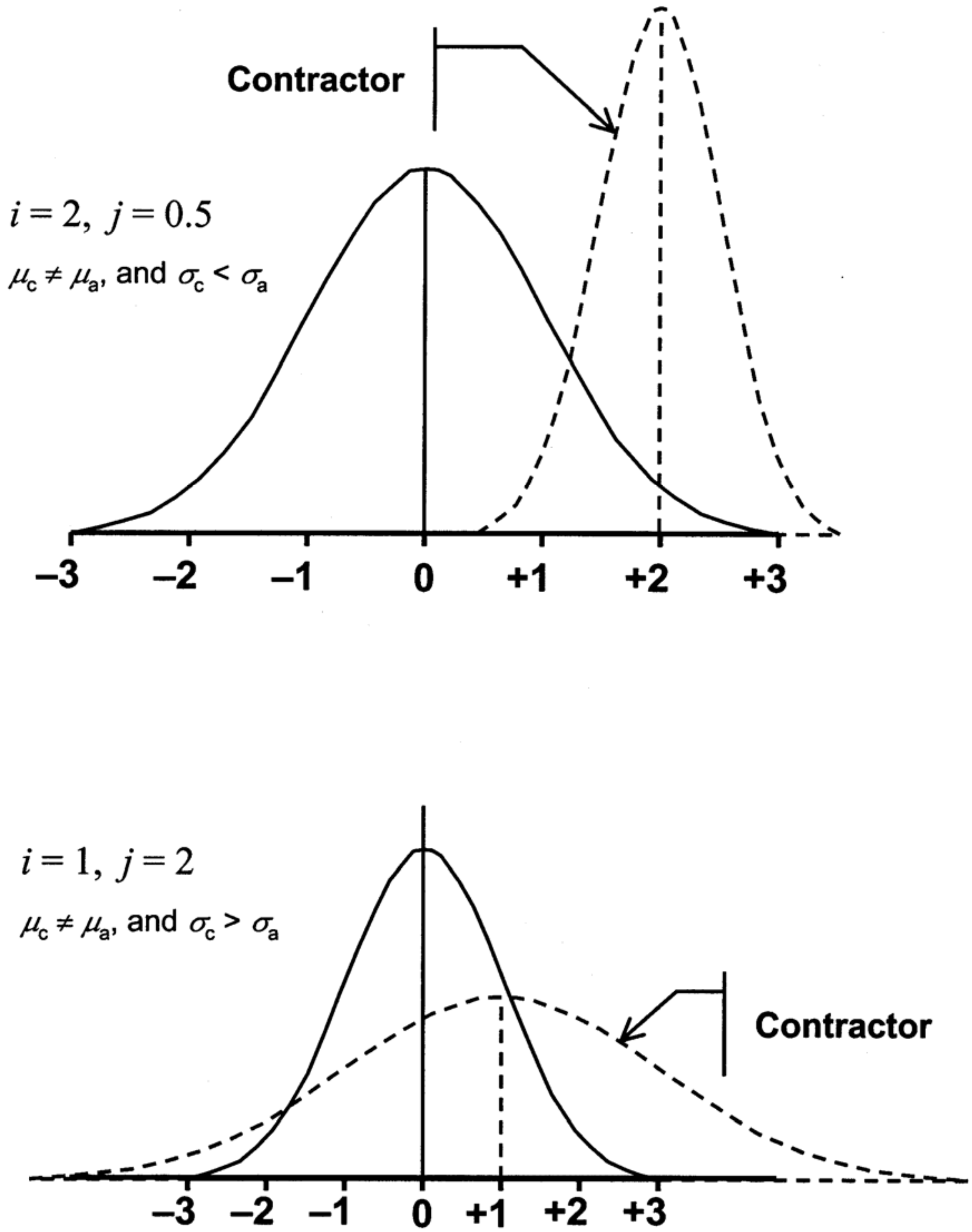


Figure 48c. Example 3 of some of the cases considered in the average run length analysis for the maximum allowable difference method.

The results of the analyses are presented in table 31 and figure 49. These values are based on 5000 simulated projects. As shown in the table, when $i = 0$ and $j = 1.0$ (meaning that the contractor's and the agency's populations are the same), the average run length is approximately 21.5 project lots. This is consistent with what would be expected. Since the limits are set at 2 standard deviations and since there is only 0.0455 chance of a value outside of 2 standard deviations, there is only 1 chance in 22 of declaring the populations to be different for this situation. It should also be noted in the table that the standard deviation values are nearly as large as the average run lengths. This means that for any individual simulated project, the run length could have varied greatly from the average. Indeed, for this case, the individual run lengths varied from 1 to more than 200.

Table 31 clearly shows that as the difference between the population means (i) increases, the average run length decreases since it is easier to detect a difference between the two populations. This is also true for the ratio of the population standard deviations (j).

Table 31. Average run length results for the single split-sample method (5000 simulated lots).

Mean Difference, units of agency's σ	Contractor's σ ÷ Agency's σ	Run Length	
		Average	Std. Dev.
0	0.5	85.57	85.44
	1.0	21.55	20.88
	1.5	8.43	8.04
	2.0	4.83	4.19
1	0.5	19.16	19.11
	1.0	9.86	9.14
	1.5	5.83	5.25
	2.0	4.07	3.53
2	0.5	4.38	3.82
	1.0	3.58	3.03
	1.5	3.10	2.56
	2.0	2.67	2.09
3	0.5	1.77	1.14
	1.0	1.85	1.27
	1.5	1.88	1.29
	2.0	1.88	1.30

Paired *t*-Test

Since the maximum allowable difference is not a very powerful test, another procedure that uses multiple test results to conduct a more powerful hypothesis test can be used. For the case in which it is desirable to compare more than one pair of split-sample test results, the *t*-test for paired measurements (i.e., the paired *t*-test) can be used. This test uses the differences between pairs of tests and determines whether the average difference is statistically different from zero. Thus, it is the difference *within* the pairs, not between the pairs, that is being tested. The *t*-statistic for the paired *t*-test is:

$$t = \frac{\left|\overline{X}_d\right|}{\frac{s_d}{\sqrt{n}}} \tag{7}$$

where: $\overline{X}_d$ = average of the differences between the split-sample test results

s_d = standard deviation of the differences between the split-sample test results

n = number of split samples

The calculated *t*-value is then compared to the critical value (t_{crit}) obtained from a table of *t*-values at a level of $\alpha/2$ and $n - 1$ degrees of freedom. Computer programs, such as Microsoft® Excel, contain statistical test procedures for the paired *t*-test. This makes the implementation process straightforward.

OC Curves: OC curves can be consulted to evaluate the performance of the paired *t*-test in identifying the differences between population means. OC curves are useful in answering the question, "How many pairs of test results should be used?" This form of the OC curve, for a given level of α, plots on the vertical axis the probability of either not detecting (β) or detecting ($1 - \beta$) a difference between two populations. The standardized difference between the two population means is plotted on the horizontal axis.

For a paired *t*-test, the standardized difference (d) is measured as:

$$d = \frac{\left|\mu_c - \mu_a\right|}{\sigma_d} \tag{8}$$

where: $\left|\mu_c - \mu_a\right|$ = true absolute difference between the mean (μ_c) of the contractor's test result population (which is unknown) and the mean (μ_a) of the agency's test result population (which is unknown)

σ_d = standard deviation of the true population of signed differences between the paired tests (which is unknown)

The OC curves are developed for a given level of significance (α). OC curves for α values of 0.05 and 0.01 are shown in figures 49 and 50, respectively. It is evident from the OC curves that for any probability of not detecting a difference (β (value on the vertical axis)), the required n will increase as the difference (d) decreases (value on the horizontal axis). In some cases, the desired β or difference may require prohibitively large sample sizes. In that case, a compromise must be made between the discriminating power desired, the cost of the amount of testing required, and the risk of claiming a difference when none exists.

To use this OC curve, the true standard deviation of the signed differences (σ_d) is assumed to be known (or approximated based on past data or published literature). After experience is gained with the process, σ_d can be more accurately defined and a better idea of the required number of tests can be determined.

As an example of how to use the OC curves, assume that the number of pairs of split-sample tests for verification of some test method is desired. The probability of *not* detecting a difference (β) is chosen as 10 percent or 0.10. (Some OC curves, which are often called *power curves*, use $1 - \beta$ (known as the power of the test) on the vertical axis; however, the only difference is the scale change (in this case, $1 - \beta$) being 90 percent or 0.90.) Assume that the absolute difference between μ_c and μ_a should not be greater than 20 units, that the standard deviation of the differences is 20 units, and that α is selected as 0.05. This produces a d value of $20 \div 20 = 1.0$. Reading this value on the horizontal axis and a β of 0.20 on the vertical axis shows that about 10 paired split-sample tests are necessary for the comparison.

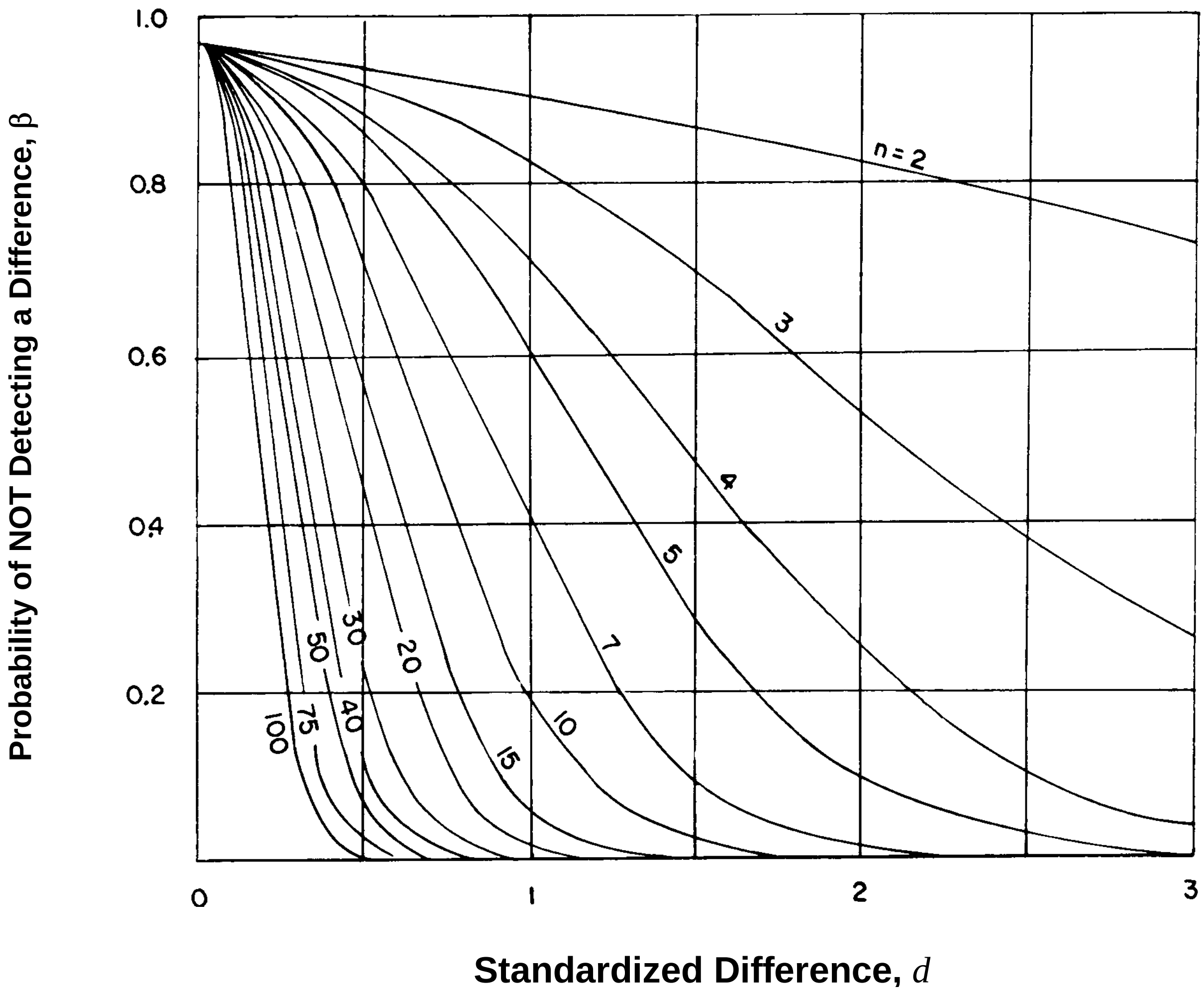


Figure 49. OC curves for a two-sided *t*-test (α = 0.05) (Natrella, M.G., "Experimental Statistics," *National Bureau of Standards Handbook 91*, 1963).

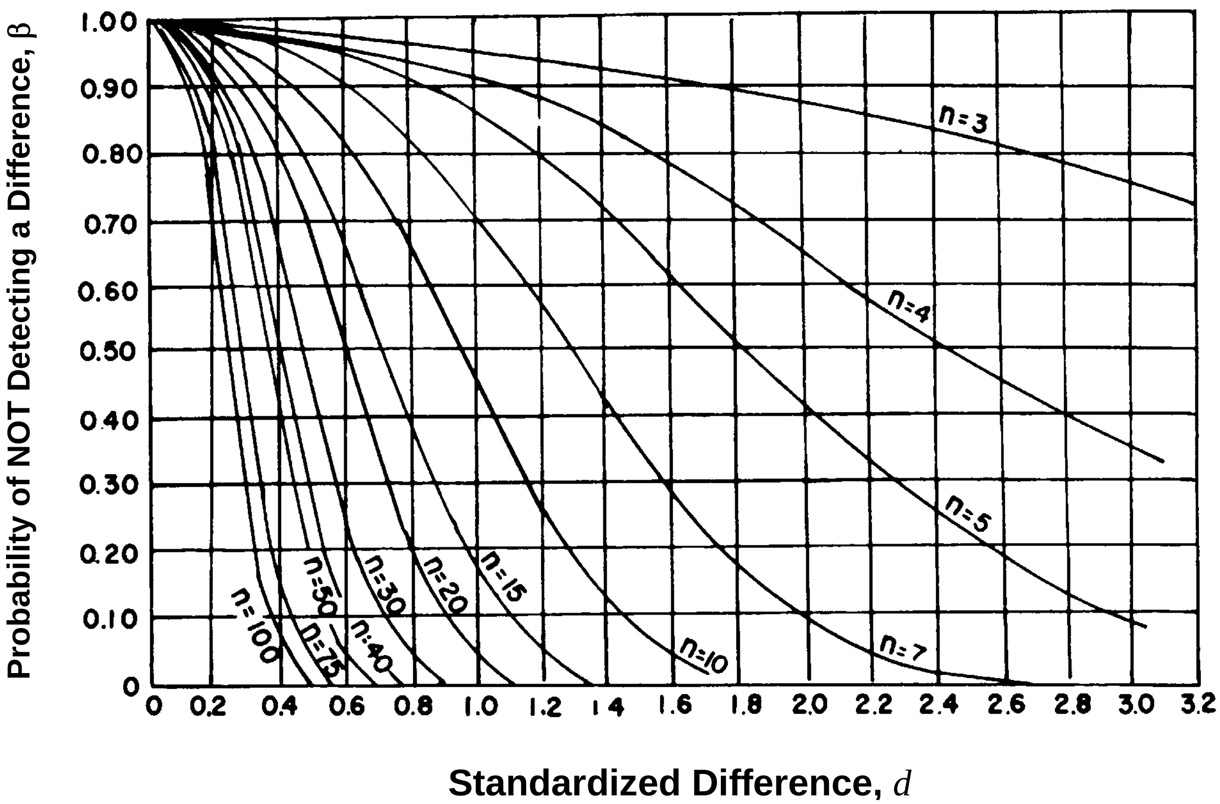


Figure 50. OC curves for a two-sided *t*-test ($\alpha = 0.01$) (Natrella, M.G., "Experimental Statistics," *National Bureau of Standards Handbook 91*, 1963).

PROCESS VERIFICATION

Procedures to verify the overall process should be based on independent samples so that all of the components of variability (i.e., process, materials, sampling, and testing) are present. Two procedures for comparing independently obtained samples appear in the *AASHTO Implementation Manual for Quality Assurance*.[2] These two methods appear in the AASHTO manual in appendix G, which is based on the comparison of a single agency test with 5 to 10 contractor tests, and in appendix H, which is based on the use of the *F*-test and *t*-test to compare a number of agency tests with a number of contractor tests. These methods are referred to as the AASHTO appendix G method and the AASHTO appendix H method, respectively. Each of these methods is discussed and analyzed in the following sections.

AASHTO Appendix G Method

In this method, a single agency test result must fall within an interval that is defined from the average and range of 5 to 10 contractor test results. The allowable interval within which the agency's test must fall is $\bar{X} \pm CR$, where $\bar{X}$ and R are the mean and range, respectively, of the contractor's tests, and C is a factor that varies with the number of contractor tests. The factor C is the product of a factor to estimate the sample standard deviation from the sample range and the t-value for the 99th percentile of the t-distribution. This is not a particularly efficient approach, although this statement can be made for any method that is based on the use of a single agency test. Table 32 indicates the allowable interval based on the number of contractor tests.

Table 32. Allowable intervals for the AASHTO appendix G method.

Number of Contractor Tests	Allowable Interval
10	$\bar{X} \pm 0.91R$
9	$\bar{X} \pm 0.97R$
8	$\bar{X} \pm 1.05R$
7	$\bar{X} \pm 1.17R$
6	$\bar{X} \pm 1.33R$
5	$\bar{X} \pm 1.61R$

OC Curves: Computer simulation was used to develop OC curves (plotted as power curves) that indicate the probability of detecting a difference between test populations with various differences in means and in the ratios of their standard deviations. The differences between the means of the contractor's and the agency's populations ($\Delta = (\mu_{Contr} - \mu_{Agency})/\sigma_{Agency}$), stated in units of the agency's standard deviation, were varied from 0 to 3.0. Various ratios of the contractor's standard deviation to the agency's standard deviation ($\sigma_{Contr}/\sigma_{Agency}$) were varied from 0.50 to 3.00.

Since there are two parameters that varied, OC surfaces were plotted, with each surface representing a different number of contractor tests (5 to 10) that were compared to a single agency test. These OC surfaces are shown in figure 51. As shown in the plots, the power of this procedure is quite low, even when a large number of contractor tests are used and when there are large differences in the means and standard deviations for the contractor's and the agency's populations. For example, for five contractor tests, even when the contractor's standard deviation is three times that of the agency and the contractor's mean is three of the agency's standard deviations from the agency's mean, there is less than a 50-percent chance of detecting a difference. Even if the number of contractor tests is 10, the probability of detecting a difference is still less than 60 percent.

Average Run Length: The method in appendix G was also evaluated based on the average run length. Various actual differences between the contractor's and the agency's population means and standard deviations were considered in the analysis. In the results that are presented, i refers to the difference (stated in units of the agency's population standard deviation) between the agency's and the contractor's population means. Also, j refers to the ratio of the contractor's population standard deviation to the agency's population standard deviation. In the analyses, i values of 0, 1, 2, and 3 were used, while j values of 0.5, 1.0, 1.5, and 2.0 were used.

The results of the simulation analyses, for the case of five contractor tests and one agency test per lot, are presented in table 33. The use of 5 and 10 contractor tests represents the upper and lower bounds, respectively, for the results since these are the fewest and most tests for the procedure. As shown in table 33, the run lengths can be quite large, particularly when the contractor's population standard deviation is larger than that of the agency. The values in the table are based on 5000 simulated projects.

Also note that the use of 10 tests gives a better performance than that of 5 tests when the contractor's standard deviation is equal to or less than that of the agency (ratios of 1.0 and 0.5). However, the opposite is true when the contractor's standard deviation is greater than that of the agency (ratios of 1.5 and 2.0). This is contrary to the desire to use a larger sample to identify the differences between the contractor's and the agency's populations.

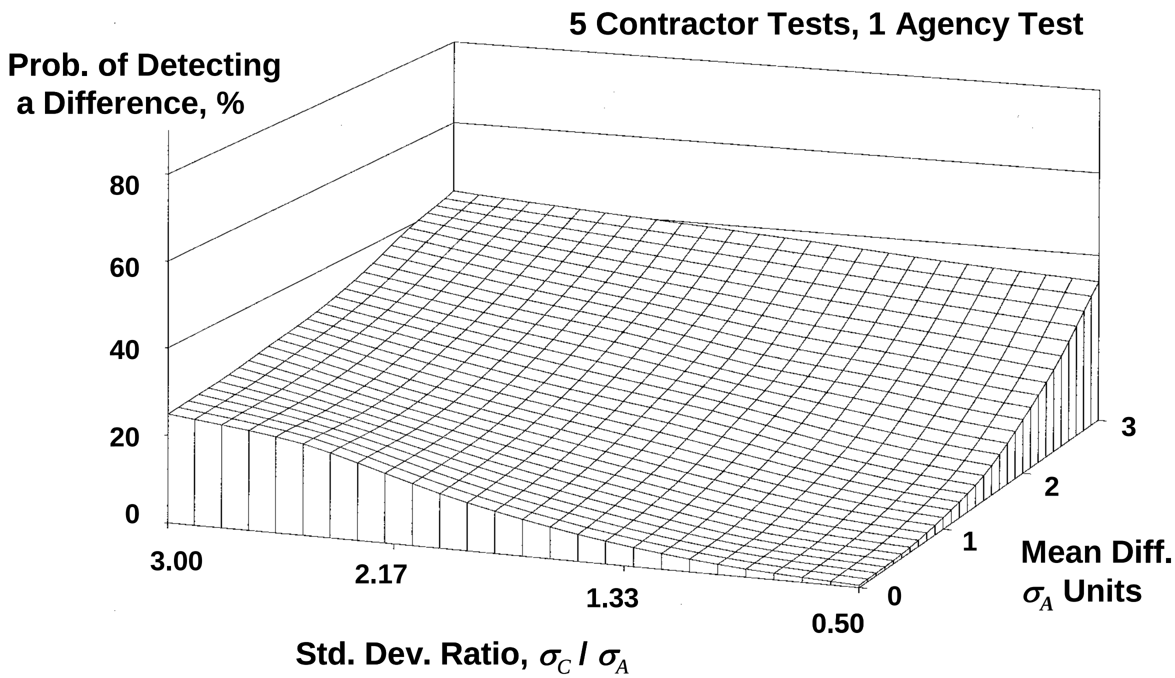


Figure 51a. OC Surfaces (also called *power surfaces*) for the appendix G method for 5 contractor tests compared to a single agency test.

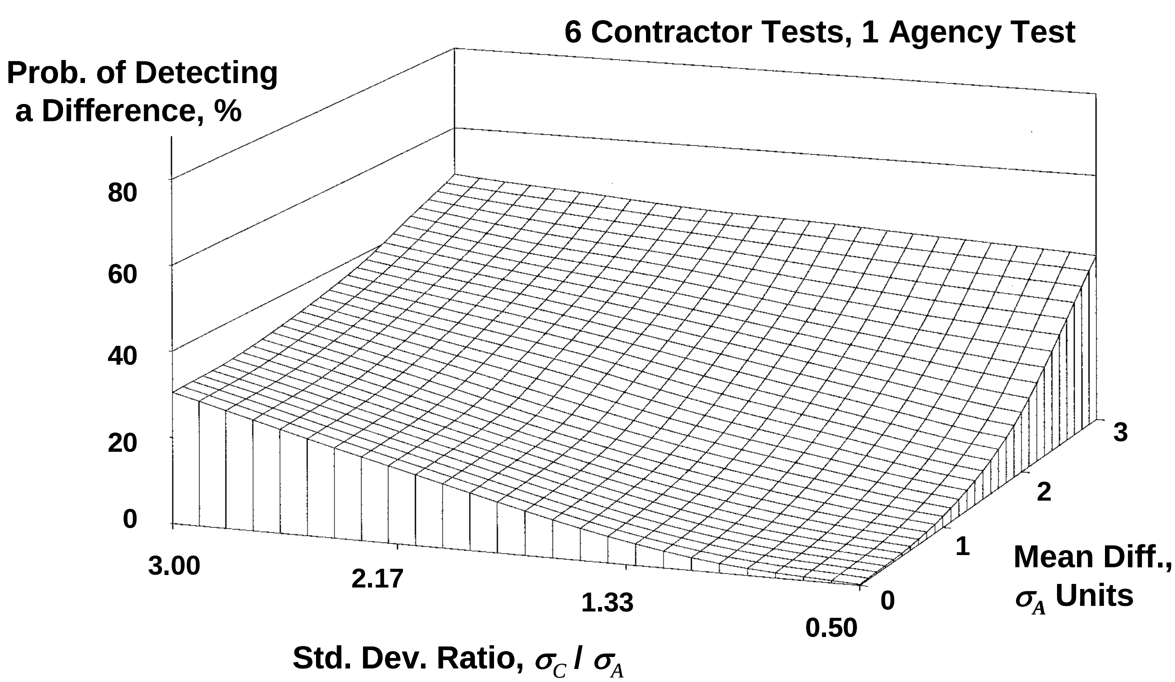


Figure 51b. OC surfaces (also called *power* surfaces) for the appendix G method for 6 contractor tests compared to a single agency test.

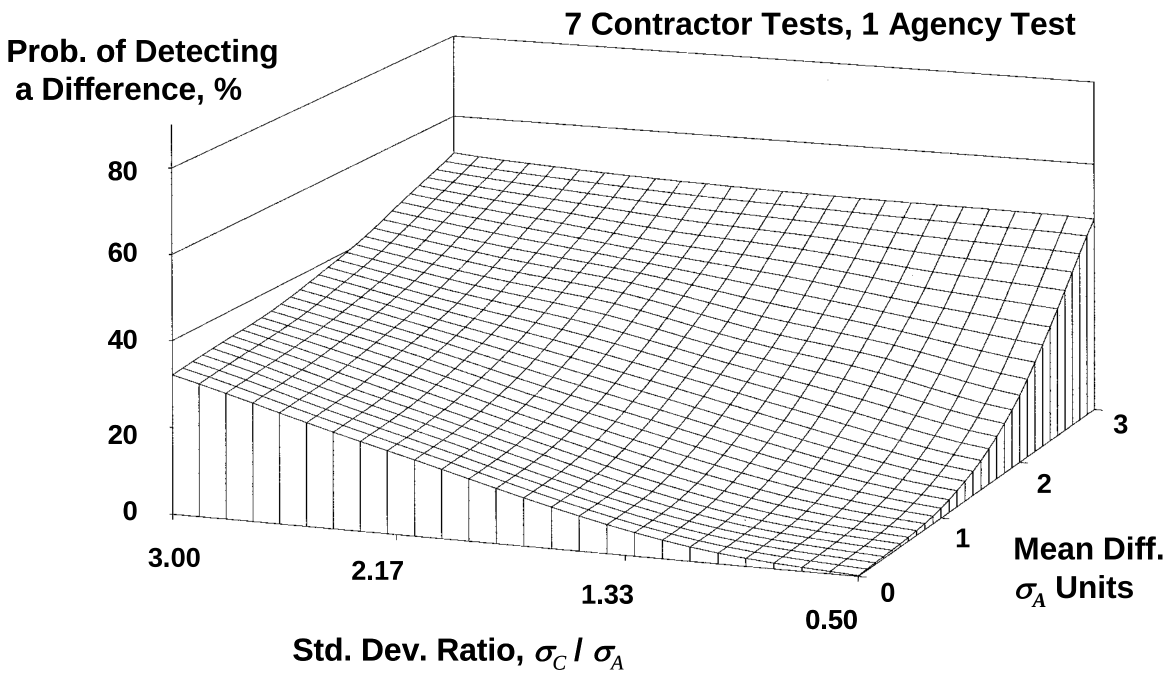


Figure 51c. OC surfaces (also called *power* surfaces) for the appendix G method for 7 contractor tests compared to a single agency test.

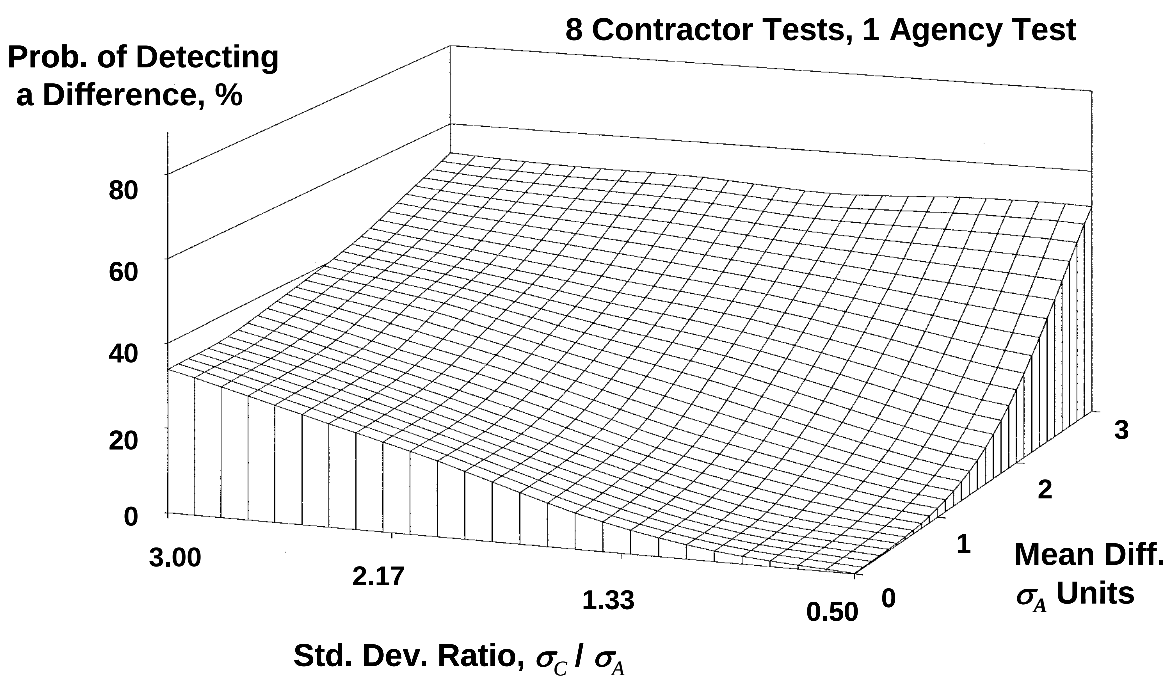


Figure 51d. OC surfaces (also called *power surfaces*) for the appendix G method for 8 contractor tests compared to a single agency test.

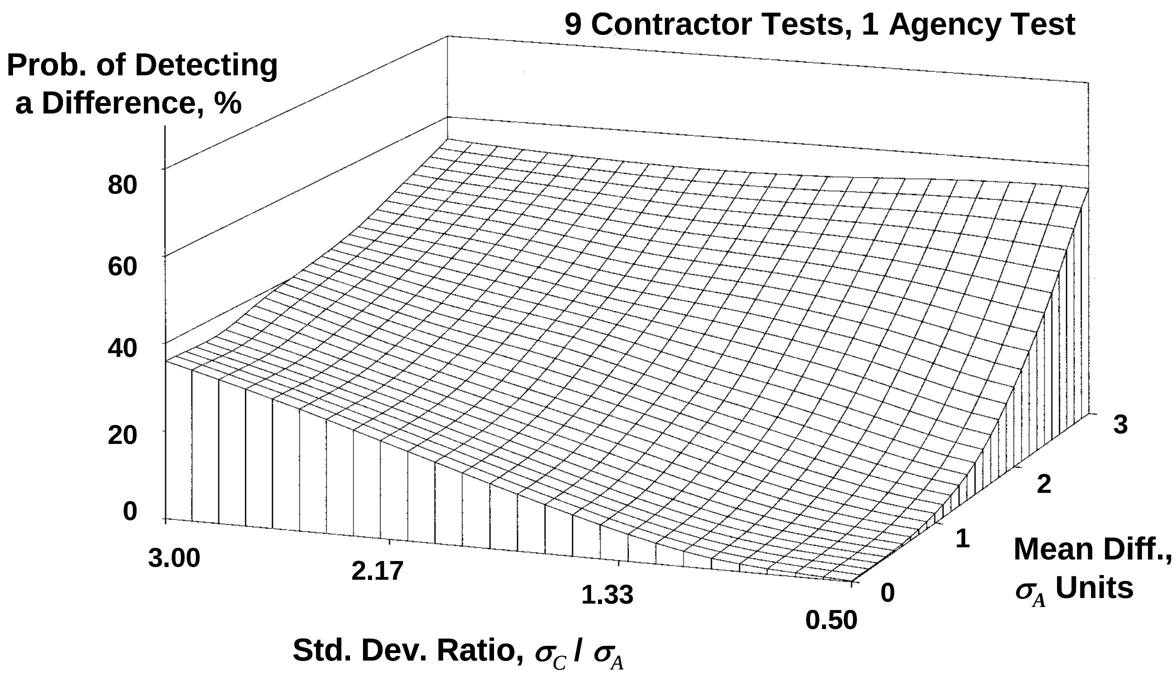


Figure 51e. OC surfaces (also called *power surfaces*) for the appendix G method for 9 contractor tests compared to a single agency test.

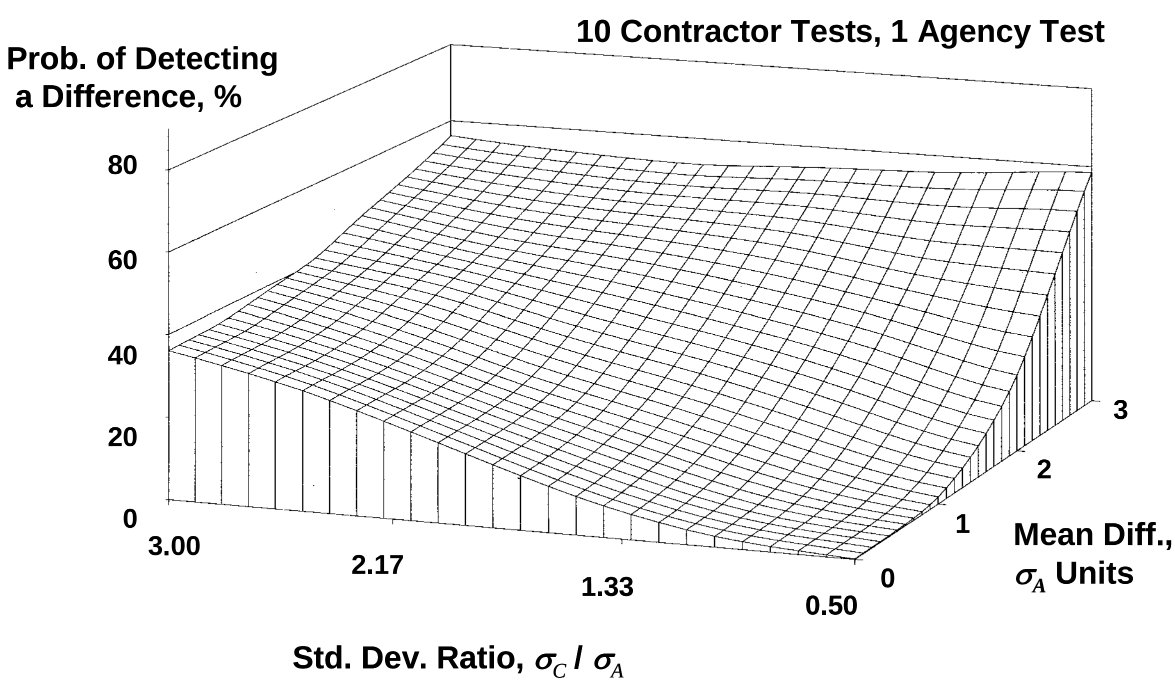


Figure 51f. OC surfaces (also called *power surfaces*) for the appendix G method for 10 contractor tests compared to a single agency test.

Table 33. Average run length results for the appendix G method (5000 simulated lots).

Mean Difference, units of agency's σ	Contractor's σ ÷ Agency's σ	Run Length	
		Average	Std. Dev.
5 Contractor Tests and 1 Agency Test			
0	0.5	7.92	7.57
	1.0	43.30	42.68
	1.5	124.19	126.40
	2.0	234.45	234.56
1	0.5	4.04	3.51
	1.0	18.04	17.78
	1.5	54.78	53.93
	2.0	114.63	114.98
2	0.5	1.82	1.24
	1.0	6.21	5.69
	1.5	17.61	17.23
	2.0	39.30	38.33
3	0.5	1.22	0.51
	1.0	2.88	2.34
	1.5	7.23	6.80
	2.0	16.23	15.74
10 Contractor Tests and 1 Agency Test			
0	0.5	5.15	4.70
	1.0	40.50	39.90
	1.5	230.83	226.93
	2.0	887.62	882.77
1	0.5	2.74	2.18
	1.0	12.76	12.04
	1.5	62.33	61.14
	2.0	229.00	227.47
2	0.5	1.39	0.73
	1.0	3.76	3.32
	1.5	13.30	12.61
	2.0	46.17	46.19
3	0.5	1.07	0.28
	1.0	1.75	1.20
	1.5	4.46	3.94
	2.0	12.77	12.15

AASHTO Appendix H Method

This procedure involves two hypothesis tests where the null hypothesis for each test is that the contractor's tests and the agency's tests are from the same population. In other words, the null hypotheses are that the variability of the two data sets is equal for the *F*-test and that the means of the two data sets are equal for the *t*-test.

The procedures for the *F*-test and the *t*-test are more complicated and involved than that for the appendix G method discussed above. The *F*-test and the *t*-test approach also requires more agency test results before a comparison can be made. However, the use of the *F*-test and the *t*-test is much more statistically sound and has more power to detect actual differences than the appendix G method, which relies on a single agency test for the comparison. Any comparison method that is based on a single test result will not be very effective in detecting differences between data sets.

When comparing two data sets that are assumed to be normally distributed, it is important to compare both the means and the variances. A different test is used for each of these comparisons. The *F*-test provides a method for comparing the variances (standard deviations squared) of the two sets of data. The differences in the means are assessed by the *t*-test. To simplify the use of these tests, they are available as built-in functions in computer spreadsheet programs such as Microsoft® Excel. For this reason, the procedures involved are not discussed in this report. The procedures are fully discussed in the QA manual that was prepared as part of this project.(1)

A question that needs to be answered is: What power do these statistical tests have, when used with small to moderate sample sizes, to declare that various differences in the means and variances are statistically significant? This question is addressed separately for the *F*-test and the *t*-test with the development of the OC curves in the following sections.

***F*-Test for Variances (Equal Sample Sizes):** Suppose that we have two sets of measurements that are assumed to come from normally distributed populations and we wish to conduct a test to see if they come from populations that have the same variances (i.e., $\sigma_x^2 = \sigma_y^2$). Furthermore, suppose that we select a level of significance of $\alpha = 0.05$, meaning that we are allowing up to a 5-percent chance of incorrectly deciding that the variances are different when they are really the same. If we assume that these two samples are $x_1, x_2, \ldots x_{nx}$ and $y_1, y_2, \ldots y_{ny}$, we can calculate the sample variances s_x^2 and s_y^2, and construct:

$$F = s_x^2 / s_y^2 \tag{9}$$

and accept $H_o : \sigma_x^2 = \sigma_y^2$ for the values of F in the interval $[F_{1-\alpha/2, nx-1, ny-1}, F_{\alpha/2, nx-1, ny-1}]$.

For this two-sided or two-tailed test, figure 52 shows the probability that we have accepted the two samples as coming from populations with the same variability. This probability is usually referred to as β and the power of the test is usually referred to as $1 - \beta$. Notice that the horizontal axis is the quantity λ, where $\lambda = \sigma_x/\sigma_y$, the true standard deviation ratio. Thus, for $\lambda = 1$, where the hypothesis of equal variance should certainly be accepted, it is accepted with a probability of

0.95, reduced from 1.00 only by the magnitude of our type I error risk (α). One significant limiting factor for the use of figure 52 is the restriction that $n_x = n_y = n$. This limitation is addressed in subsequent sections of the report.

Example: Suppose that we have $n_x = 6$ contractor tests and $n_y = 6$ agency tests, conduct an $\alpha =$ 0.05 level test and accept (or fail to reject) that these two sets of tests represent populations with equal variances. What power did our test have to discern whether the populations from which these two sets of tests came were really rather different in variability? Suppose that the true population standard deviation of the contractor's tests (σ_x) was twice as large as that of the agency's tests (σ_y), giving $\lambda = 2$. If we enter figure 52 with $\lambda = 2$ and $n_x = n_y = 6$, we find that $\beta \approx 0.74$ or that the power $(1 - \beta)$ is 0.26. This tells us that with samples of $n_x = 6$ and $n_y = 6$, we only have a 26-percent chance of detecting a standard deviation ratio of 2 (and, correspondingly, a fourfold difference in variance) as being different.

Suppose that we are not comfortable with the power of 0.26, so subsequently we increase the number of tests used. Then suppose that we now have $n_x = 20$ and $n_y = 20$. If we again consider $\lambda = 2$, we can determine from figure 52 that the power of detecting these sets of tests as coming from populations with unequal variances to be more than 0.80 (approximately 82 to 83 percent). If we proceed to conduct our *F*-test with these two samples and conclude that the underlying variances are equal, we will certainly feel much more comfortable with our conclusions.

Figure 53 gives the appropriate OC curves to be used if we choose to conduct an $\alpha = 0.01$ level test. Again, we see that for equal variances σ_x^2 and σ_y^2 (i.e., $\lambda = 1$), that $\beta = 0.99$, reduced from 1.00 only by the size of α.

***F*-Test for Variances (Unequal Sample Sizes):** Up to now, the discussions and OC curves have been limited to equal sample sizes. Routines were developed for this project to calculate the power for this test for any combination of sample sizes n_x and n_y. There are obviously an infinite number of possible combinations for n_x and n_y. Thus, it is not possible to present OC curves for every possibility. However, three sets of tables were developed to provide a subset of power calculations using some sample sizes that are of potential interest for comparing the contractor's and the agency's samples. These power calculations are presented in table form since there are too many variables to be presented in a single chart, and the data can be presented in a more compact form in tables than in a long series of charts. Table 34 gives power values for all combinations of sample sizes of 3 to 10, with the ratio of the two subpopulation standard deviations = 1, 2, 3, 4, and 5. Table 35 gives power values for the same sample sizes, but with the standard deviation ratios = 0.0, 0.2, 0.4, 0.6, 0.8, and 1.0. Table 36 gives power values for all combinations for sample sizes = 5, 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 90, and 100, with the standard deviation ratio = 1, 2, or 3.

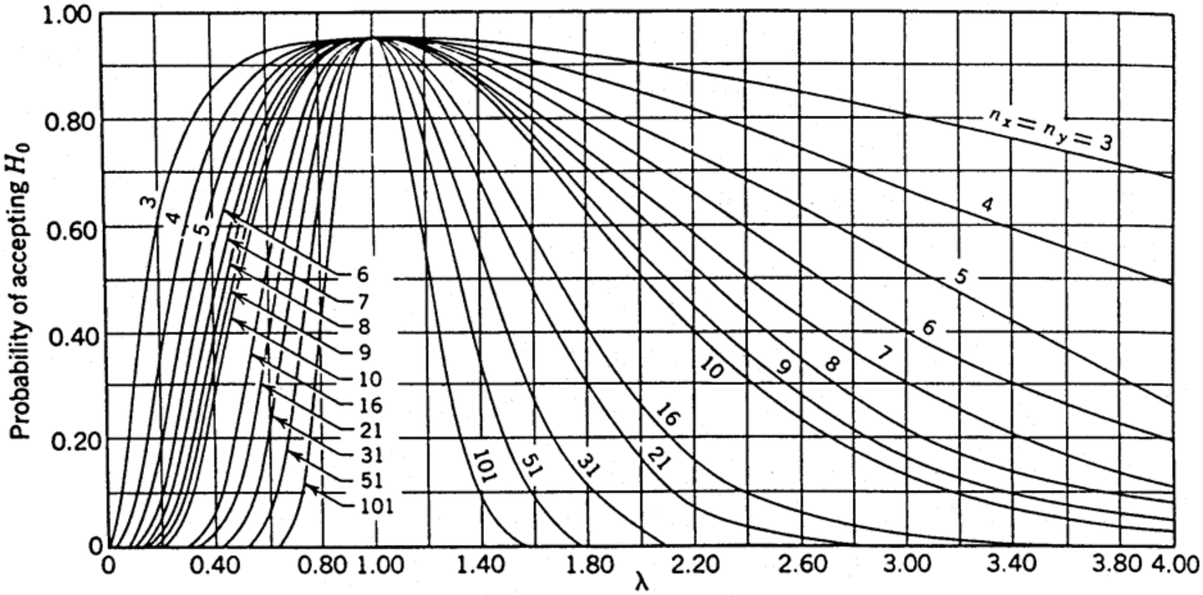


Figure 52. OC curves for the two-sided *F*-test for level of significance α = 0.05 (Bowker, A.H., and G.J. Lieberman, *Engineering Statistics*).

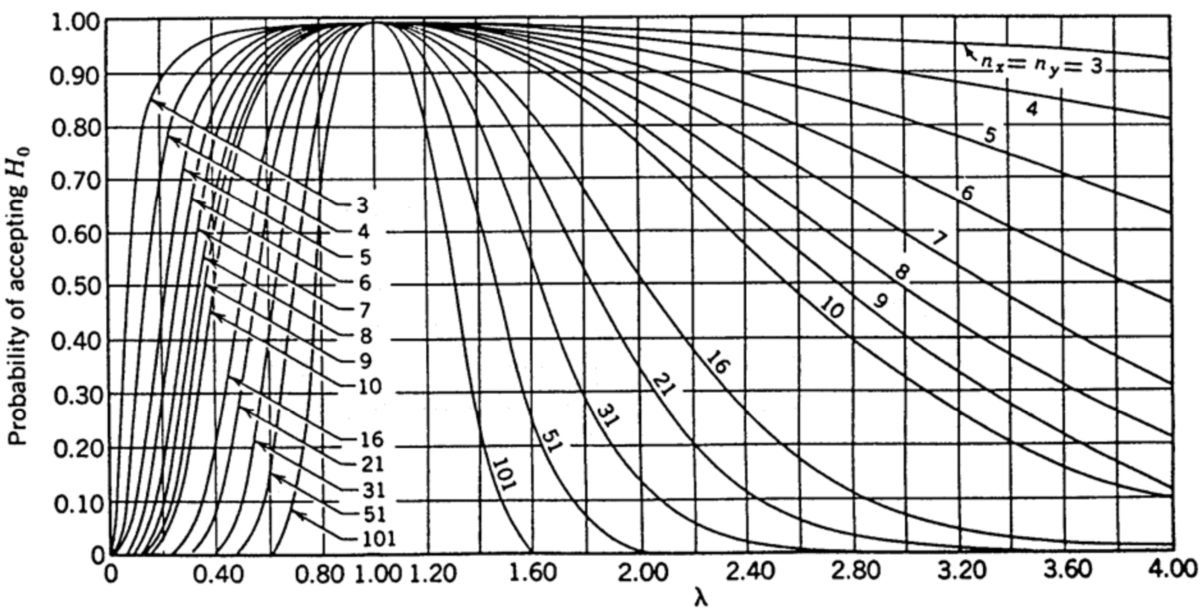


Figure 53. OC curves for the two-sided *F*-test for level of significance α = 0.01 (Bowker, A.H., and G.J. Lieberman, *Engineering Statistics*).

Table 34. *F*-test power values for n = 3-10 and s-ratio λ = 1-5.

λ	n_x	n_y	Power
1	3	3	0.05000
		4	0.05000
		5	0.05000
		6	0.05000
		7	0.05000
		8	0.05000
		9	0.05000
		10	0.05000
	4	3	0.05000
		4	0.05000
		5	0.05000
		6	0.05000
		7	0.05000
		8	0.05000
		9	0.05000
		10	0.05000
	5	3	0.05000
		4	0.05000
		5	0.05000
		6	0.05000
		7	0.05000
		8	0.05000
		9	0.05000
		10	0.05000
	6	3	0.05000
		4	0.05000
		5	0.05000
		6	0.05000
		7	0.05000
		8	0.05000
		9	0.05000
		10	0.05000
	7	3	0.05000
		4	0.05000
		5	0.05000
		6	0.05000
		7	0.05000
		8	0.05000
		9	0.05000
		10	0.05000
	8	3	0.05000
		4	0.05000
		5	0.05000
		6	0.05000
		7	0.05000
		8	0.05000
		9	0.05000
		10	0.05000
1	9	3	0.05000
		4	0.05000
		5	0.05000
		6	0.05000
		7	0.05000
		8	0.05000
		9	0.05000
		10	0.05000
	10	3	0.05000
		4	0.05000
		5	0.05000
		6	0.05000
		7	0.05000
		8	0.05000
		9	0.05000
		10	0.05000
2	3	3	0.09939
		4	0.09753
		5	0.09663
		6	0.09620
		7	0.09600
		8	0.09590
		9	0.09586
		10	0.09585
	4	3	0.14835
		4	0.15169
		5	0.15385
		6	0.15544
		7	0.15668
		8	0.15767
		9	0.15848
		10	0.15915
	5	3	0.19036
		4	0.20240
		5	0.21041
		6	0.21622
		7	0.22064
		8	0.22413
		9	0.22694
		10	0.22926
	6	3	0.22309
		4	0.24464
		5	0.25968
		6	0.27093
		7	0.27968
		8	0.28669
		9	0.29243
		10	0.29722
2	7	3	0.24820
		4	0.27854
		5	0.30055
		6	0.31744
		7	0.33086
		8	0.34179
		9	0.35087
		10	0.35853
	8	3	0.26768
		4	0.30567
		5	0.33401
		6	0.35619
		7	0.37410
		8	0.38888
		9	0.40129
		10	0.41187
	9	3	0.28308
		4	0.32758
		5	0.36144
		6	0.38837
		7	0.41036
		8	0.42869
		9	0.44421
		10	0.45752
	10	3	0.29549
		4	0.34549
		5	0.38414
		6	0.41521
		7	0.44081
		8	0.46230
		9	0.48060
		10	0.49639
3	3	3	0.19034
		4	0.19354
		5	0.19556
		6	0.19696
		7	0.19798
		8	0.19875
		9	0.19934
		10	0.19981
	4	3	0.31171
		4	0.33525
		5	0.35007
		6	0.36030
		7	0.36777
		8	0.37347
		9	0.37795
		10	0.38157

Table 34. *F*-test power values for *n* = 3-10 and s-ratio λ = 1-5 (continued).

λ	n_x	n_y	Power
3	5	3	0.39758
		4	0.44454
		5	0.47603
		6	0.49872
		7	0.51588
		8	0.52931
		9	0.54011
		10	0.54899
	6	3	0.45403
		4	0.51906
		5	0.56396
		6	0.59696
		7	0.62225
		8	0.64225
		9	0.65846
		10	0.67186
	7	3	0.49230
		4	0.57007
		5	0.62436
		6	0.66443
		7	0.69516
		8	0.71943
		9	0.73906
		10	0.75523
	8	3	0.51945
		4	0.60623
		5	0.66693
		6	0.71159
		7	0.74565
		8	0.77236
		9	0.79378
		10	0.81129
	9	3	0.53955
		4	0.63285
		5	0.69797
		6	0.74560
		7	0.78161
		8	0.80958
		9	0.83177
		10	0.84970
	10	3	0.55494
		4	0.65311
		5	0.72136
		6	0.77092
		7	0.80803
		8	0.83654
		9	0.85890
		10	0.87675
4	3	3	0.29251
		4	0.30367
		5	0.31010
		6	0.31427
		7	0.31717
		8	0.31930
		9	0.32093
		10	0.32222
	4	3	0.46558
		4	0.51179
		5	0.54104
		6	0.56126
		7	0.57608
		8	0.58742
		9	0.59637
		10	0.60363
	5	3	0.56455
		4	0.63665
		5	0.68356
		6	0.71649
		7	0.74084
		8	0.75955
		9	0.77437
		10	0.78638
	6	3	0.62143
		4	0.70759
		5	0.76314
		6	0.80150
		7	0.82932
		8	0.85027
		9	0.86652
		10	0.87943
	7	3	0.65697
		4	0.75074
		5	0.81002
		6	0.84993
		7	0.87808
		8	0.89866
		9	0.91416
		10	0.92613
	8	3	0.68090
		4	0.77901
		5	0.83976
		6	0.87961
		7	0.90692
		8	0.92628
		9	0.94042
		10	0.95100
4	9	3	0.69798
		4	0.79871
		5	0.85988
		6	0.89907
		7	0.92520
		8	0.94321
		9	0.95598
		10	0.96525
	10	3	0.71073
		4	0.81311
		5	0.87423
		6	0.91256
		7	0.93751
		8	0.95427
		9	0.96583
		10	0.97399
5	3	3	0.39165
		4	0.41270
		5	0.42481
		6	0.43266
		7	0.43815
		8	0.44219
		9	0.44530
		10	0.44776
	4	3	0.58713
		4	0.64932
		5	0.68814
		6	0.71467
		7	0.73394
		8	0.74858
		9	0.76007
		10	0.76932
	5	3	0.68068
		4	0.76196
		5	0.81171
		6	0.84479
		7	0.86811
		8	0.88527
		9	0.89836
		10	0.90860
	6	3	0.72975
		4	0.81790
		5	0.86956
		6	0.90223
		7	0.92409
		8	0.93936
		9	0.95041
		10	0.95864

Table 34. *F*-test power values for *n* = 3-10 and s-ratio λ = 1-5 (continued).

λ	n_x	n_y	Power
5	7	3	0.75893
		4	0.84940
		5	0.90024
		6	0.93086
		7	0.95030
		8	0.96318
		9	0.97201
		10	0.97824
	8	3	0.77800
		4	0.86909
		5	0.91845
		6	0.94695
		7	0.96423
		8	0.97513
		9	0.98225
		10	0.98704
	9	3	0.79133
		4	0.88238
		5	0.93024
		6	0.95690
		7	0.97244
		8	0.98184
		9	0.98772
		10	0.99150
	10	3	0.80115
		4	0.89188
		5	0.93838
		6	0.96351
		7	0.97767
		8	0.98594
		9	0.99092
		10	0.99400

Table 35. *F*-test power values for *n* = 3-10 and s-ratio λ = 0-1.

λ	n_x	n_y	Power
0.0	3	3	1.00000
		4	1.00000
		5	1.00000
		6	1.00000
		7	1.00000
		8	1.00000
		9	1.00000
		10	1.00000
	4	3	1.00000
		4	1.00000
		5	1.00000
		6	1.00000
		7	1.00000
		8	1.00000
		9	1.00000
		10	1.00000
	5	3	1.00000
		4	1.00000
		5	1.00000
		6	1.00000
		7	1.00000
		8	1.00000
		9	1.00000
		10	1.00000
	6	3	1.00000
		4	1.00000
		5	1.00000
		6	1.00000
		7	1.00000
		8	1.00000
		9	1.00000
		10	1.00000
	7	3	1.00000
		4	1.00000
		5	1.00000
		6	1.00000
		7	1.00000
		8	1.00000
		9	1.00000
		10	1.00000
	8	3	1.00000
		4	1.00000
		5	1.00000
		6	1.00000
		7	1.00000
		8	1.00000
		9	1.00000
		10	1.00000

λ	n_x	n_y	Power
0.0	9	3	1.00000
		4	1.00000
		5	1.00000
		6	1.00000
		7	1.00000
		8	1.00000
		9	1.00000
		10	1.00000
	10	3	1.00000
		4	1.00000
		5	1.00000
		6	1.00000
		7	1.00000
		8	1.00000
		9	1.00000
		10	1.00000
0.2	3	3	0.39165
		4	0.58713
		5	0.68068
		6	0.72975
		7	0.75893
		8	0.77800
		9	0.79133
		10	0.80115
	4	3	0.41270
		4	0.64932
		5	0.76196
		6	0.81790
		7	0.84940
		8	0.86909
		9	0.88238
		10	0.89188
	5	3	0.42481
		4	0.68814
		5	0.81171
		6	0.86956
		7	0.90024
		8	0.91845
		9	0.93024
		10	0.93838
	6	3	0.43266
		4	0.71467
		5	0.84479
		6	0.90223
		7	0.93086
		8	0.94695
		9	0.95690
		10	0.96351

λ	n_x	n_y	Power
0.2	7	3	0.43815
		4	0.73394
		5	0.86811
		6	0.92409
		7	0.95030
		8	0.96423
		9	0.97244
		10	0.97767
	8	3	0.44219
		4	0.74858
		5	0.88527
		6	0.93936
		7	0.96318
		8	0.97513
		9	0.98184
		10	0.98594
	9	3	0.44530
		4	0.76007
		5	0.89836
		6	0.95041
		7	0.97201
		8	0.98225
		9	0.98772
		10	0.99092
	10	3	0.44776
		4	0.76932
		5	0.90860
		6	0.95864
		7	0.97824
		8	0.98704
		9	0.99150
		10	0.99400
0.4	3	3	0.14221
		4	0.22806
		5	0.29564
		6	0.34398
		7	0.37868
		8	0.40429
		9	0.42380
		10	0.43906
	4	3	0.14250
		4	0.24034
		5	0.32488
		6	0.38884
		7	0.43614
		8	0.47159
		9	0.49879
		10	0.52015

Table 35. *F*-test power values for *n* = 3-10 and s-ratio λ = 0-1 (continued).

λ	n_x	n_y	Power
0.4	5	3	0.14291
		4	0.24808
		5	0.34448
		6	0.42028
		7	0.47749
		8	0.52079
		9	0.55411
		10	0.58029
	6	3	0.14332
		4	0.25345
		5	0.35863
		6	0.44371
		7	0.50889
		8	0.55851
		9	0.59674
		10	0.62671
	7	3	0.14369
		4	0.25739
		5	0.36934
		6	0.46187
		7	0.53357
		8	0.58837
		9	0.63057
		10	0.66355
	8	3	0.14399
		4	0.26041
		5	0.37772
		6	0.47638
		7	0.55351
		8	0.61261
		9	0.65804
		10	0.69341
	9	3	0.14424
		4	0.26278
		5	0.38447
		6	0.48825
		7	0.56996
		8	0.63266
		9	0.68076
		10	0.71805
	10	3	0.14445
		4	0.26470
		5	0.39001
		6	0.49813
		7	0.58375
		8	0.64952
		9	0.69984
		10	0.73868

λ	n_x	n_y	Power
0.6	3	3	0.07564
		4	0.10273
		5	0.12665
		6	0.14614
		7	0.16173
		8	0.17425
		9	0.18444
		10	0.19283
	4	3	0.07283
		4	0.10212
		5	0.13003
		6	0.15430
		7	0.17470
		8	0.19170
		9	0.20593
		10	0.21791
	5	3	0.07120
		4	0.10174
		5	0.13222
		6	0.15988
		7	0.18396
		8	0.20461
		9	0.22225
		10	0.23736
	6	3	0.07022
		4	0.10157
		5	0.13386
		6	0.16407
		7	0.19107
		8	0.21472
		9	0.23528
		10	0.25314
	7	3	0.06960
		4	0.10153
		5	0.13516
		6	0.16736
		7	0.19675
		8	0.22292
		9	0.24600
		10	0.26628
	8	3	0.06919
		4	0.10155
		5	0.13622
		6	0.17003
		7	0.20139
		8	0.22972
		9	0.25499
		10	0.27741

λ	n_x	n_y	Power
0.6	9	3	0.06891
		4	0.10161
		5	0.13711
		6	0.17223
		7	0.20526
		8	0.23545
		9	0.26265
		10	0.28698
	10	3	0.06870
		4	0.10168
		5	0.13786
		6	0.17409
		7	0.20854
		8	0.24035
		9	0.26925
		10	0.29529
0.8	3	3	0.05467
		4	0.06163
		5	0.06758
		6	0.07248
		7	0.07649
		8	0.07980
		9	0.08255
		10	0.08487
	4	3	0.05202
		4	0.05929
		5	0.06587
		6	0.07156
		7	0.07642
		8	0.08057
		9	0.08412
		10	0.08719
	5	3	0.05017
		4	0.05755
		5	0.06448
		6	0.07067
		7	0.07612
		8	0.08090
		9	0.08508
		10	0.08875
	6	3	0.04883
		4	0.05626
		5	0.06340
		6	0.06995
		7	0.07584
		8	0.08109
		9	0.08577
		10	0.08994

Table 35. *F*-test power values for *n* = 3-10 and s-ratio λ = 0-1 (continued).

λ	n_x	n_y	Power
0.8	7	3	0.04785
		4	0.05529
		5	0.06258
		6	0.06938
		7	0.07560
		8	0.08124
		9	0.08633
		10	0.09092
	8	3	0.04709
		4	0.05453
		5	0.06193
		6	0.06893
		7	0.07541
		8	0.08136
		9	0.08680
		10	0.09175
	9	3	0.04650
		4	0.05393
		5	0.06141
		6	0.06856
		7	0.07527
		8	0.08148
		9	0.08721
		10	0.09248
	10	3	0.04603
		4	0.05345
		5	0.06099
		6	0.06827
		7	0.07516
		8	0.08159
		9	0.08757
		10	0.09312
1.0	3	3	0.05000
		4	0.05000
		5	0.05000
		6	0.05000
		7	0.05000
		8	0.05000
		9	0.05000
		10	0.05000
	4	3	0.05000
		4	0.05000
		5	0.05000
		6	0.05000
		7	0.05000
		8	0.05000
		9	0.05000
		10	0.05000
1.0	5	3	0.05000
		4	0.05000
		5	0.05000
		6	0.05000
		7	0.05000
		8	0.05000
		9	0.05000
		10	0.05000
	6	3	0.05000
		4	0.05000
		5	0.05000
		6	0.05000
		7	0.05000
		8	0.05000
		9	0.05000
		10	0.05000
	7	3	0.05000
		4	0.05000
		5	0.05000
		6	0.05000
		7	0.05000
		8	0.05000
		9	0.05000
		10	0.05000
	8	3	0.05000
		4	0.05000
		5	0.05000
		6	0.05000
		7	0.05000
		8	0.05000
		9	0.05000
		10	0.05000
	9	3	0.05000
		4	0.05000
		5	0.05000
		6	0.05000
		7	0.05000
		8	0.05000
		9	0.05000
		10	0.05000
	10	3	0.05000
		4	0.05000
		5	0.05000
		6	0.05000
		7	0.05000
		8	0.05000
		9	0.05000
		10	0.05000

Table 36. *F*-test power values for *n* = 5-100 and s-ratio λ = 1-3.

λ	n_x	n_y	Power
		5	0.05
		10	0.05
		15	0.05
		20	0.05
		25	0.05
		30	0.05
	5	40	0.05
		50	0.05
		60	0.05
		70	0.05
		80	0.05
		90	0.05
		100	0.05
		5	0.05
		10	0.05
		15	0.05
		20	0.05
		25	0.05
		30	0.05
1	10	40	0.05
		50	0.05
		60	0.05
		70	0.05
		80	0.05
		90	0.05
		100	0.05
		5	0.05
		10	0.05
		15	0.05
		20	0.05
		25	0.05
		30	0.05
	15	40	0.05
		50	0.05
		60	0.05
		70	0.05
		80	0.05
		90	0.05
		100	0.05

λ	n_x	n_y	Power
		5	0.05
		10	0.05
		15	0.05
		20	0.05
		25	0.05
		30	0.05
	20	40	0.05
		50	0.05
		60	0.05
		70	0.05
		80	0.05
		90	0.05
		100	0.05
		5	0.05
		10	0.05
		15	0.05
		20	0.05
		25	0.05
		30	0.05
1	25	40	0.05
		50	0.05
		60	0.05
		70	0.05
		80	0.05
		90	0.05
		100	0.05
		5	0.05
		10	0.05
		15	0.05
		20	0.05
		25	0.05
		30	0.05
	30	40	0.05
		50	0.05
		60	0.05
		70	0.05
		80	0.05
		90	0.05
		100	0.05

λ	n_x	n_y	Power
		5	0.05
		10	0.05
		15	0.05
		20	0.05
		25	0.05
		30	0.05
	40	40	0.05
		50	0.05
		60	0.05
		70	0.05
		80	0.05
		90	0.05
		100	0.05
		5	0.05
		10	0.05
		15	0.05
		20	0.05
		25	0.05
		30	0.05
1	50	40	0.05
		50	0.05
		60	0.05
		70	0.05
		80	0.05
		90	0.05
		100	0.05
		5	0.05
		10	0.05
		15	0.05
		20	0.05
		25	0.05
		30	0.05
	60	40	0.05
		50	0.05
		60	0.05
		70	0.05
		80	0.05
		90	0.05
		100	0.05

Table 36. *F*-test power values for *n* = 5-100 and s-ratio λ = 1-3 (continued).

λ	n_x	n_y	Power
1	70	5	0.05
		10	0.05
		15	0.05
		20	0.05
		25	0.05
		30	0.05
		40	0.05
		50	0.05
		60	0.05
		70	0.05
		80	0.05
		90	0.05
		100	0.05
	80	5	0.05
		10	0.05
		15	0.05
		20	0.05
		25	0.05
		30	0.05
		40	0.05
		50	0.05
		60	0.05
		70	0.05
		80	0.05
		90	0.05
		100	0.05
	90	5	0.05
		10	0.05
		15	0.05
		20	0.05
		25	0.05
		30	0.05
		40	0.05
		50	0.05
		60	0.05
		70	0.05
		80	0.05
		90	0.05
		100	0.05

λ	n_x	n_y	Power
1	100	5	0.05
		10	0.05
		15	0.05
		20	0.05
		25	0.05
		30	0.05
		40	0.05
		50	0.05
		60	0.05
		70	0.05
		80	0.05
		90	0.05
		100	0.05
2	5	5	0.21041
		10	0.22926
		15	0.23658
		20	0.24043
		25	0.24281
		30	0.24442
		40	0.24646
		50	0.24770
		60	0.24853
		70	0.24913
		80	0.24958
		90	0.24993
		100	0.25022
	10	5	0.38414
		10	0.49639
		15	0.55109
		20	0.58353
		25	0.60501
		30	0.62027
		40	0.64053
		50	0.65336
		60	0.66221
		70	0.66869
		80	0.67363
		90	0.67753
		100	0.68068

λ	n_x	n_y	Power
2	15	5	0.45487
		10	0.62152
		15	0.70573
		20	0.75560
		25	0.78820
		30	0.81099
		40	0.84054
		50	0.85870
		60	0.87092
		70	0.87969
		80	0.88626
		90	0.89137
		100	0.89545
	20	5	0.49087
		10	0.68548
		15	0.78230
		20	0.83747
		25	0.87192
		30	0.89495
		40	0.92304
		50	0.93906
		60	0.94918
		70	0.95606
		80	0.96099
		90	0.96468
		100	0.96753
	25	5	0.51241
		10	0.72299
		15	0.82516
		20	0.88085
		25	0.91389
		30	0.93485
		40	0.95864
		50	0.97099
		60	0.97817
		70	0.98272
		80	0.98578
		90	0.98795
		100	0.98955

Table 36. *F*-test power values for *n* = 5-100 and s-ratio λ = 1-3 (continued).

λ	n_x	n_y	Power
2	30	5	0.52669
		10	0.74730
		15	0.85174
		20	0.90637
		25	0.93725
		30	0.95585
		40	0.97551
		50	0.98476
		60	0.98968
		70	0.99256
		80	0.99436
		90	0.99556
		100	0.99639
	40	5	0.54439
		10	0.77664
		15	0.88220
		20	0.93379
		25	0.96067
		30	0.97548
		40	0.98924
		50	0.99462
		60	0.99702
		70	0.99821
		80	0.99886
		90	0.99923
		100	0.99945
	50	5	0.55491
		10	0.79358
		15	0.89881
		20	0.94770
		25	0.97160
		30	0.98387
		40	0.99414
		50	0.99757
		60	0.99888
		70	0.99943
		80	0.99969
		90	0.99982
		100	0.99989
2	60	5	0.56187
		10	0.80456
		15	0.90914
		20	0.95588
		25	0.97764
		30	0.98820
		40	0.99632
		50	0.99869
		60	0.99948
		70	0.99977
		80	0.99989
		90	0.99995
		100	0.99997
	70	5	0.56683
		10	0.81224
		15	0.91614
		20	0.96120
		25	0.98137
		30	0.99073
		40	0.99745
		50	0.99921
		60	0.99972
		70	0.99989
		80	0.99996
		90	0.99998
		100	0.99999
	80	5	0.57053
		10	0.81791
		15	0.92118
		20	0.96490
		25	0.98387
		30	0.99235
		40	0.99810
		50	0.99947
		60	0.99984
		70	0.99994
		80	0.99998
		90	0.99999
		100	1.00000
2	90	5	0.57339
		10	0.82226
		15	0.92497
		20	0.96762
		25	0.98564
		30	0.99345
		40	0.99851
		50	0.99962
		60	0.99989
		70	0.99997
		80	0.99999
		90	1.00000
		100	1.00000
	100	5	0.57568
		10	0.82571
		15	0.92793
		20	0.96968
		25	0.98696
		30	0.99425
		40	0.99879
		50	0.99972
		60	0.99993
		70	0.99998
		80	0.99999
		90	1.00000
		100	1.00000
3	5	5	0.47603
		10	0.54899
		15	0.57700
		20	0.59187
		25	0.60108
		30	0.60736
		40	0.61537
		50	0.62026
		60	0.62355
		70	0.62593
		80	0.62772
		90	0.62911
		100	0.63024

Table 36. *F*-test power values for *n* = 5-100 and s-ratio λ = 1-3 (continued).

λ	n_x	n_y	Power
3	10	5	0.72136
		10	0.87675
		15	0.92836
		20	0.95158
		25	0.96404
		30	0.97154
		40	0.97985
		50	0.98420
		60	0.98681
		70	0.98853
		80	0.98973
		90	0.99062
		100	0.99130
	15	5	0.78336
		10	0.93786
		15	0.97640
		20	0.98918
		25	0.99431
		30	0.99669
		40	0.99860
		50	0.99928
		60	0.99957
		70	0.99972
		80	0.99980
		90	0.99985
		100	0.99988
	20	5	0.80975
		10	0.95808
		15	0.98816
		20	0.99597
		25	0.99841
		30	0.99930
		40	0.99982
		50	0.99994
		60	0.99998
		70	0.99999
		80	0.99999
		90	1.00000
		100	1.00000

λ	n_x	n_y	Power
3	25	5	0.82417
		10	0.96743
		15	0.99254
		20	0.99797
		25	0.99936
		30	0.99977
		40	0.99996
		50	0.99999
		60	1.00000
		70	1.00000
		80	1.00000
		90	1.00000
		100	1.00000
	30	5	0.83321
		10	0.97267
		15	0.99463
		20	0.99877
		25	0.99968
		30	0.99990
		40	0.99999
		50	1.00000
		60	1.00000
		70	1.00000
		80	1.00000
		90	1.00000
		100	1.00000
	40	5	0.84390
		10	0.97822
		15	0.99654
		20	0.99938
		25	0.99987
		30	0.99997
		40	1.00000
		50	1.00000
		60	1.00000
		70	1.00000
		80	1.00000
		90	1.00000
		100	1.00000

λ	n_x	n_y	Power
3	50	5	0.84999
		10	0.98107
		15	0.99738
		20	0.99960
		25	0.99993
		30	0.99999
		40	1.00000
		50	1.00000
		60	1.00000
		70	1.00000
		80	1.00000
		90	1.00000
		100	1.00000
	60	5	0.85393
		10	0.98279
		15	0.99783
		20	0.99971
		25	0.99996
		30	0.99999
		40	1.00000
		50	1.00000
		60	1.00000
		70	1.00000
		80	1.00000
		90	1.00000
		100	1.00000
	70	5	0.85668
		10	0.98394
		15	0.99812
		20	0.99976
		25	0.99997
		30	1.00000
		40	1.00000
		50	1.00000
		60	1.00000
		70	1.00000
		80	1.00000
		90	1.00000
		100	1.00000

Table 36. *F*-test power values for *n* = 5-100 and s-ratio λ = 1-3 (continued).

λ	n_x	n_y	Power
3	80	5	0.85871
		10	0.98476
		15	0.99831
		20	0.99980
		25	0.99998
		30	1.00000
		40	1.00000
		50	1.00000
		60	1.00000
		70	1.00000
		80	1.00000
		90	1.00000
		100	1.00000
	90	5	0.86026
		10	0.98537
		15	0.99844
		20	0.99983
		25	0.99998
		30	1.00000
		40	1.00000
		50	1.00000
		60	1.00000
		70	1.00000
		80	1.00000
		90	1.00000
		100	1.00000
	100	5	0.86150
		10	0.98584
		15	0.99855
		20	0.99985
		25	0.99998
		30	1.00000
		40	1.00000
		50	1.00000
		60	1.00000
		70	1.00000
		80	1.00000
		90	1.00000
		100	1.00000

From these tables, it is obvious that the limiting factor in how well the *F*-test will be able to identify differences will be the number of agency verification tests. The power of the *F*-test is limited not by the larger of the sample sizes, but by the smaller of the sample sizes. For example, in table 34, when $n_x = 3$ and $n_y = 10$, the power is only about 20 percent, even when there is a threefold difference in the true standard deviations (i.e., $\lambda = 3$). The limiting aspect of the smaller sample size is also noticeable in table 36 for larger sample sizes. For example, for $\lambda = 2$ and for $n_y = 100$, the power when $n_x = 5$ is only about 25 percent. The power increases to 68 percent for $n_x = 10$, 90 percent for $n_x = 15$, and 97 percent for $n_x = 20$. Since the agency will have fewer verification tests than the number of contractor tests, the agency's verification sampling and testing rate will determine the power to identify variability differences when they exist.

***t*-Test for Means:** As with the appendix G method, the performance of the *t*-test for means can be evaluated with OC curves or by considering the average run length.

OC Curves: Suppose that we have two sets of measurements that are assumed to be from normally distributed populations and that we wish to conduct a two-sided or two-tailed test to see if these populations have equal means (i.e., $\mu_x = \mu_y$). Suppose that we assume that these two samples are from populations with unknown, but equal, variances. If these two samples are $x_1, x_2, \ldots, x_{nx}$, with sample mean $\overline{X}$ and sample variance s_x^2, and $y_1, y_2, \ldots, y_{ny}$, with sample mean $\overline{Y}$ and sample variance s_y^2, we can calculate:

$$t = \frac{\overline{X} - \overline{Y}}{\sqrt{\frac{s_x^2(n_x - 1) + s_y^2(n_y - 1)}{n_x + n_y - 2}} \times \sqrt{\frac{1}{n_x} + \frac{1}{n_y}}} \tag{10}$$

and accept H_0: $\mu_x = \mu_y$ for values of *t* in the interval $[-t_{\alpha/2, n_x + n_y - 2}, t_{\alpha/2, n_x + n_y - 2}]$.

For this test, figure 49 or 50, depending on the α value, shows the probability that we have accepted the two samples as coming from populations with the same means. The horizontal axis scale is:

$$d = \frac{|\mu_x - \mu_y|}{\sigma} \tag{11}$$

where: $\sigma = \sigma_x = \sigma_y$ = true common population standard deviation

We can access the OC curves in figure 49 or 50 with a value for *d* of d^* and a value for *n* of n'

where:

$$n' = n_x + n_y - 1 \tag{12}$$

and

$$d^* = \frac{d}{\sqrt{n'}} \times \sqrt{\frac{n_x \times n_y}{n_x + n_y}} \tag{13}$$

Example: Suppose that we have $n_x = 8$ contractor tests and $n_y = 8$ agency tests, conduct an $\alpha =$ 0.05 level test and accept that these two sets of tests represent populations with equal means. What power did our test really have to discern if the populations from which these two sets of tests came had different means? Suppose that we consider a difference in these population means of 2 or more standard deviations as a noteworthy difference that we would like to detect with high probability. This would indicate that we are interested in $d = 2$. Calculating

$$n' = n_x + n_y - 1 = 8 + 8 - 1 = 15 \tag{14}$$

and

$$d^* = \frac{d}{\sqrt{n'}} \times \sqrt{\frac{n_x \times n_y}{n_x + n_y}} = \frac{2}{\sqrt{15}} \times \sqrt{\frac{8 \times 8}{8 + 8}} = 1.0328 \tag{15}$$

we find from figure 50 that $\beta \approx 0.05$, so that our power of detecting a mean difference of 2 or more standard deviations would be approximately 95 percent.

Now suppose that we consider an application where we still have a total of 16 tests, but with $n_x =$ 12 contractor tests and $n_y = 4$ agency tests. Suppose that we are again interested in the *t*-test performance in detecting a means difference of 2 standard deviations. Again, calculating

$$n' = n_x + n_y - 1 = 8 + 8 - 1 = 15 \tag{16}$$

but now

$$d^* = \frac{d}{\sqrt{n'}} \times \sqrt{\frac{n_x \times n_y}{n_x + n_y}} = \frac{2}{\sqrt{15}} \times \sqrt{\frac{12 \times 4}{12 + 4}} = 0.8944 \tag{17}$$

we find from figure 50 that $\beta \approx 0.12$, indicating that our power of detecting a mean difference of 2 or more standard deviations would be approximately 88 percent.

Figure 51 gives the appropriate OC curves for use in conducting an $\alpha = 0.01$ level test on the means. This figure is accessed in the same manner as described above for figure 50.

Average Run Length: The effectiveness of the *t*-test procedure was evaluated by determining the average run length in terms of project lots. The evaluation was performed by simulating 1000 projects and determining, on average, how many lots it took to determine that there was a difference between the contractor's and the agency's population means.

The results of the simulation analyses, for the case of five contractor tests and one agency test per lot, are presented in table 37. The results are shown only for the case where five contractor tests and one agency test are performed on each project lot. Similar results were obtained for cases where fewer and more contractor tests were conducted per lot. As shown in table 37, when there is no difference between the population means, the run lengths are quite large (as they should be). The values with asterisks are biased on the low side, because to speed up the simulation time, the maximum run lengths were limited to 100. Therefore, the actual average run length would be greater than those shown in the table since the maximum cutoff value was reached in more than half of the 1000 projects simulated for each *i* and *j* combination.

The average run lengths become relatively small as the actual difference between the contractor's and the agency's population means increases. This is obviously what is desired.

Table 37. Average run length results for the appendix H method (5 contractor tests and 1 agency test per lot) for 1000 simulated lots.

Mean Difference, units of agency's σ	Contractor's σ ÷ Agency's σ	Run Length	
		Average	Std. Dev.
0	0.5	55.47*	46.01*
	1.0	70.15*	41.91*
	1.5	77.78*	36.95*
	2.0	75.72*	38.56*
1	0.5	4.83	4.05
	1.0	5.75	4.28
	1.5	8.63	5.70
	2.0	9.83	5.94
2	0.5	2.60	1.18
	1.0	2.64	1.02
	1.5	3.51	1.52
	2.0	4.40	2.03
3	0.5	2.35	0.73
	1.0	2.10	0.37
	1.5	2.36	0.66
	2.0	2.88	1.03

* These values are lower than the actual values. To reduce the simulation processing time, the maximum number of lots was limited to 100. For these cases, more than half of the projects were truncated at 100 lots.

CONCLUSIONS AND RECOMMENDATIONS

Based on the analyses that were conducted and were summarized in this chapter, the following recommendations were made:

Recommendation for Test Method Verification

The comparison of a single split sample by using the maximum allowable limits (such as the D2S limits) is simple and can be done for each split sample that is obtained. However, since it is based on comparing only single data values, it is not very powerful for identifying differences where they exist. It is recommended that each individual split sample be compared using the maximum allowable limits, but that the paired *t*-test also be used on the accumulated split-sample results to allow for a comparison with more discerning power. If either of these comparisons indicates a difference, then an investigation to identify the cause of the difference should be initiated.

Recommendation for Process Verification

Since they are both based on five contractor tests and one agency test per lot, the results in tables 33 and 37 can be used to compare the appendix H and appendix G methods. The average run lengths for the appendix H method (*t*-test) were better than those for the appendix G method (single agency test compared to five contractor tests). Compared to the appendix G method, the appendix H method had longer average run lengths where there was no difference in the means and shorter lengths where there was a difference in the means. This is what is desirable in the verification procedure. The appendix H method is recommended for use in verifying the contractor's test results when the agency obtains independent samples for evaluating the total process.

From the OC curves that were developed, it is apparent that the number of agency verification tests will be the deciding factor when determining the validity of the contractor's overall process. When using the OC curves in figure 50 or 51, the lower the value of d^*, the lower the power of the test for a given number of test results. The value for d^* will decrease as the agency's portion of the total number of tests declines (this is shown in equation 13). If, in the expression under the square root sign, the total number of tests ($n_x + n_y$) is fixed, then the value of d^* will decrease as the value of either n_x or n_y goes down.

An example will illustrate this point. Suppose that the total of $n_x + n_y$ is fixed at 16, then the maximum value under the square root sign will be when $n_x = n_y = 8$. This is true because the denominator is fixed at 16 and $8 \times 8 = 64$ is larger than any other combination of numbers that total 16. As one of the values gets smaller (and the other gets correspondingly larger), the product of the two numbers will decrease, thereby decreasing d^* and reducing the power of the test.

The amount of verification sampling and testing is a subjective decision for each individual agency. However, with the OC (or power) curves and tables in this chapter, an agency can determine the risks that are associated with any frequency of verification testing and can make an informed decision regarding this testing frequency.

When using the appendix H method, first, an *F*-test is used to determine whether or not the variances (and, hence, standard deviations) are different for the two populations. The result of the *F*-test determines how the subsequent *t*-test is conducted to compare the averages of the contractor's and the agency's test results. Given some of the low powers associated with small sample sizes in tables 34 through 36, it could be argued that an agency will rarely be able to conclude from the *F*-test that a difference in variances exists. Given this fact, it may be reasonable to just assume that the populations have equal variances and run the *t*-test for equal variances and ignore the *F*-test altogether. This argument has some merit. However, with the ease of conducting the *F*-test and the *t*-test by computer, once the test results are input, there is essentially no additional effort associated with conducting the *F*-test before the *t*-test.

8. QUALITY MEASURES AND PAYMENT

CANDIDATE QUALITY MEASURES

In chapter 5, candidate quality measures were investigated with respect to their accuracy and precision when estimating the true population parameters. The candidate quality measures that were analyzed were PWL (or its complement, PD), AAD, and CI.

From the simulation analyses summarized in chapter 5, both the sample PWL and the sample AAD were unbiased estimators of the corresponding population values. They both showed decreasing variability in the estimated values as the sample size increased. For PWL, the variability values peaked at a population PWL of 50 and then decreased as the population PWL approached either 100 or 0. The AAD variability values were lowest when the population mean was centered at the target and they increased as the offset from the target of the population mean increased, leveling off as the offset approached 3 standard deviations.

The simulation analyses for CI showed similar variability relationships as were found for AAD. However, the CI appeared to be a slightly biased estimator of the population CI since all of the simulated average CI sample values were negative.

Since there appears to be no appreciable benefit to using CI rather than AAD (i.e., they have similar variability traits, but AAD is unbiased while the CI is slightly biased), the CI was eliminated from further analyses. This chapter presents the results of the further analyses that were conducted on PWL and AAD.

ANALYSES PERFORMED

The quality measure that is selected needs to be capable of being used to determine the payment factor. Therefore, it must be easy to develop an equation that relates the quality measure to the anticipated performance of the pavement. For the analyses to be conducted at this point, it was sufficient to accept the fact that the quality of the material in the population would increase as either its PWL value increased or its AAD value decreased. The relationship of payment to projected performance is discussed in a later chapter.

For each quality measure, it was necessary to determine whether or not it could provide an unbiased estimate for the payment factor for a given population. The amount of variability in the payment factor estimates for individual lots also required study. These results were both generated by computer simulation studies.

Since it is rare that only a single quality characteristic is used to determine the payment factor for a lot, it was necessary to consider how the quality measures performed when multiple acceptance characteristics were used and to consider methods for combining these individual estimated quality measures into a combined (or composite) payment factor. While it was not within the scope of the original project, it was decided that, if multiple quality characteristics were to be investigated, it would be necessary to consider potential correlations among these characteristics.

Finally, it was necessary to investigate the risks associated with the quality measures and how they would be used for acceptance decisions. The concept of risks for acceptance is similar to that discussed in chapter 7 for verification testing. If pass and fail are the only options, the acceptance decision is simply a form of hypothesis testing and the concept is fairly straightforward. When acceptance at payments other than 100 percent is possible (either incentives or disincentives), the evaluation of the risks becomes much more complicated and not nearly as clear-cut as for pass or fail decisions. The analysis of the payment risks associated with PWL and AAD quality measures was conducted for single quality characteristics and for two quality characteristics that were either independent or correlated.

PAYMENT FOR A SINGLE QUALITY CHARACTERISTIC

In chapter 5, it was determined that both PWL and AAD were unbiased estimators for the true population values. As such, they should also provide unbiased payment factor estimates when used in a payment equation. Computer simulation studies were conducted to verify that this was true.

PWL Payment Factors

Computer simulation was used to generate expected payment (EP) curves for a payment equation that based payment on the estimated PWL value. Since PWL is an unbiased estimator of the actual PWL value, the EP curve should follow the payment equation exactly. This will be true provided that an appropriate incentive provision is included in the specification. This is discussed at length in a later section on evaluating risks.

For the purposes of the simulation, any payment equation could have been used. For the simulation studies, the payment equation from the *AASHTO Quality Assurance Guide Specification* was used.[3] This payment equation is:

$$\text{Pay} = 55 + (0.5 \times \text{PWL}) \tag{18}$$

This equation was used with sample sizes = 3, 5, and 10. Figures 54 and 55 show the results of the analyses. The plots in figure 54 show the EP curves and the limits within which 95 percent of the estimated payment values fell. As expected, the EP (i.e., the average payment in the long run) follows the same line as the payment equation (equation 18). However, there is considerable variability in the payment values for individual lots. This variability (or spread about the EP line) decreases as the sample size increases from 3 to 10. This was also expected since it follows exactly what was found when PWL values were simulated and since it is the PWL values that determine the payment factor.

To further illustrate the distribution of individual lot payment factors, figure 55 presents histograms of both the individual PWL and the payment factors for populations with an actual PWL = 90 and 50, for sample sizes = 3, 5, and 10.

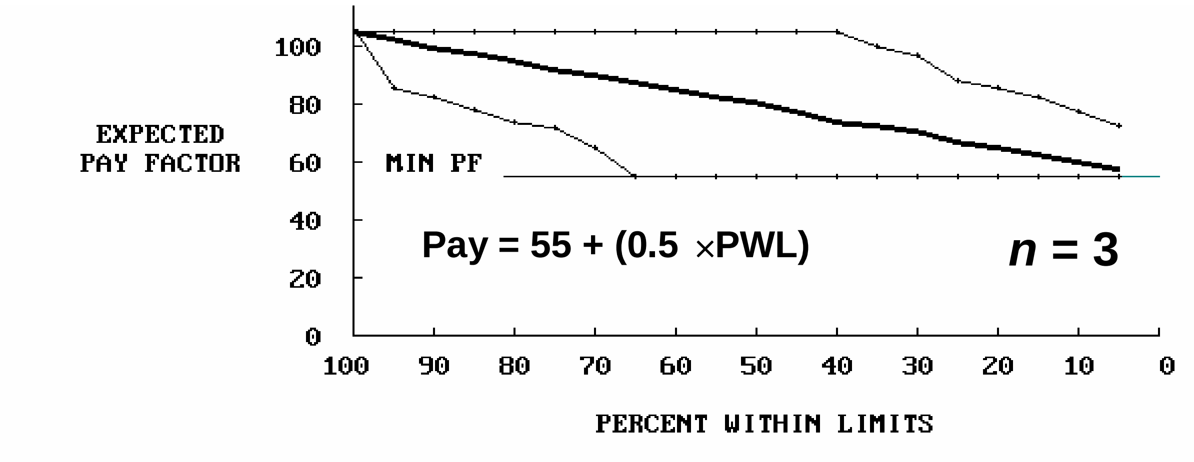


Figure 54a. EP curves for PWL payment schedule with sample size = 3.

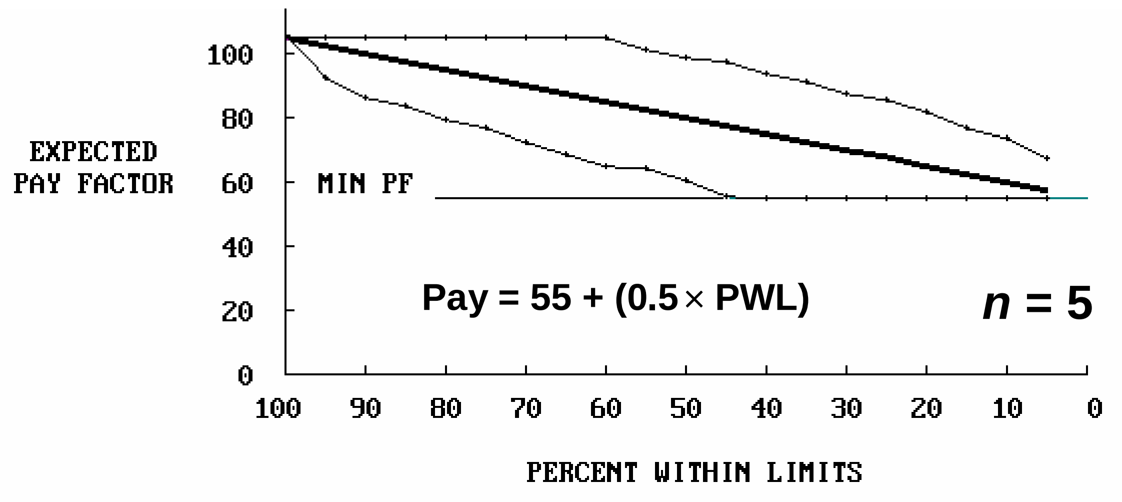


Figure 54b. EP curves for PWL payment schedule with sample size = 5.

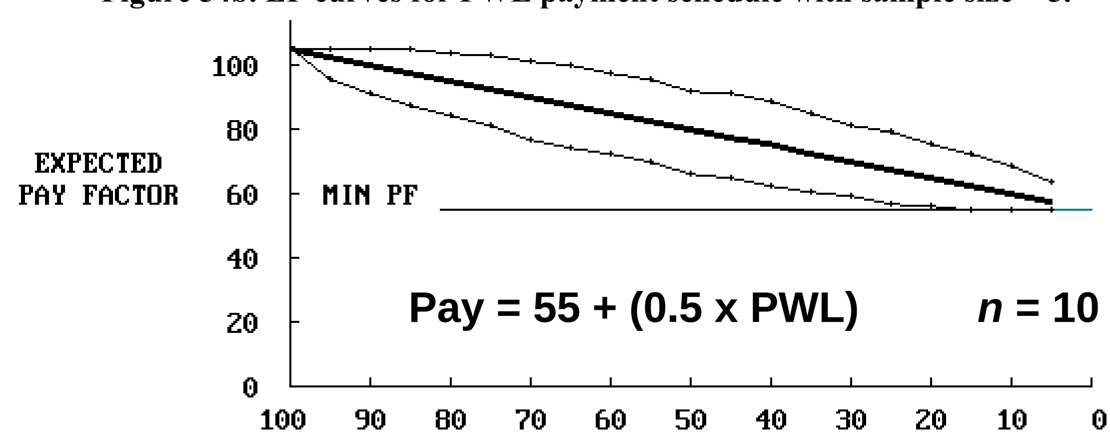


Figure 54c. EP curves for PWL payment schedule with sample size = 10.

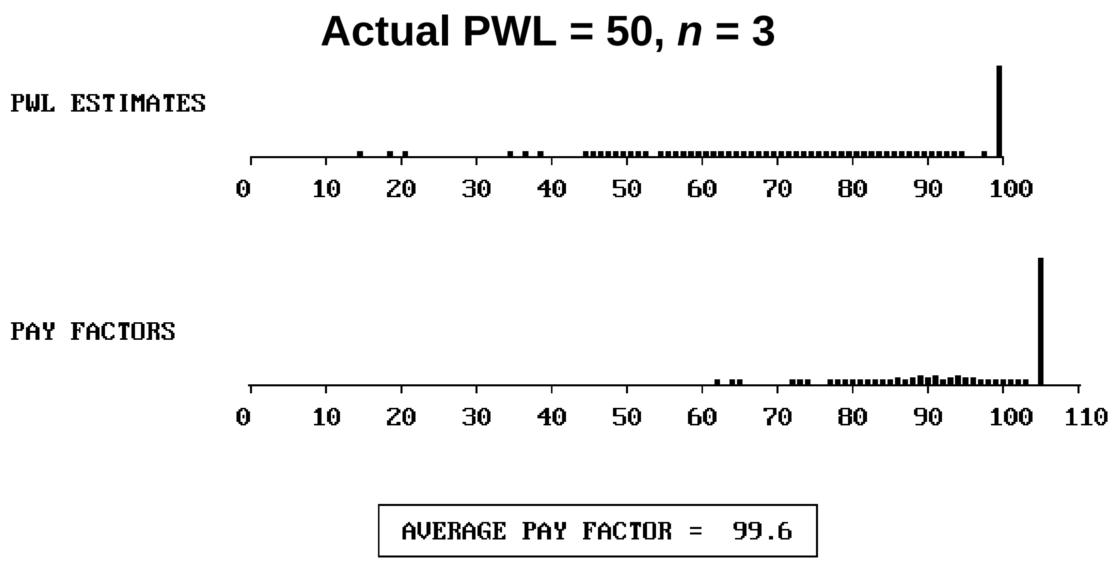


Figure 55a. Distribution of actual PWL = 90, sample size = 3, and resulting payment factors.

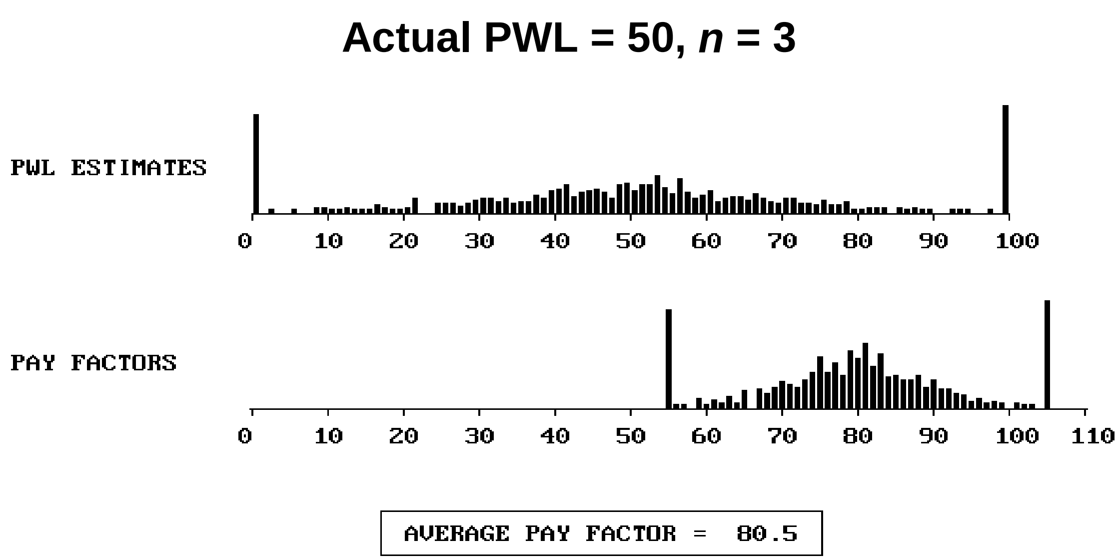


Figure 55b. Distribution of actual PWL = 50, sample size = 3, and the resulting payment factors.

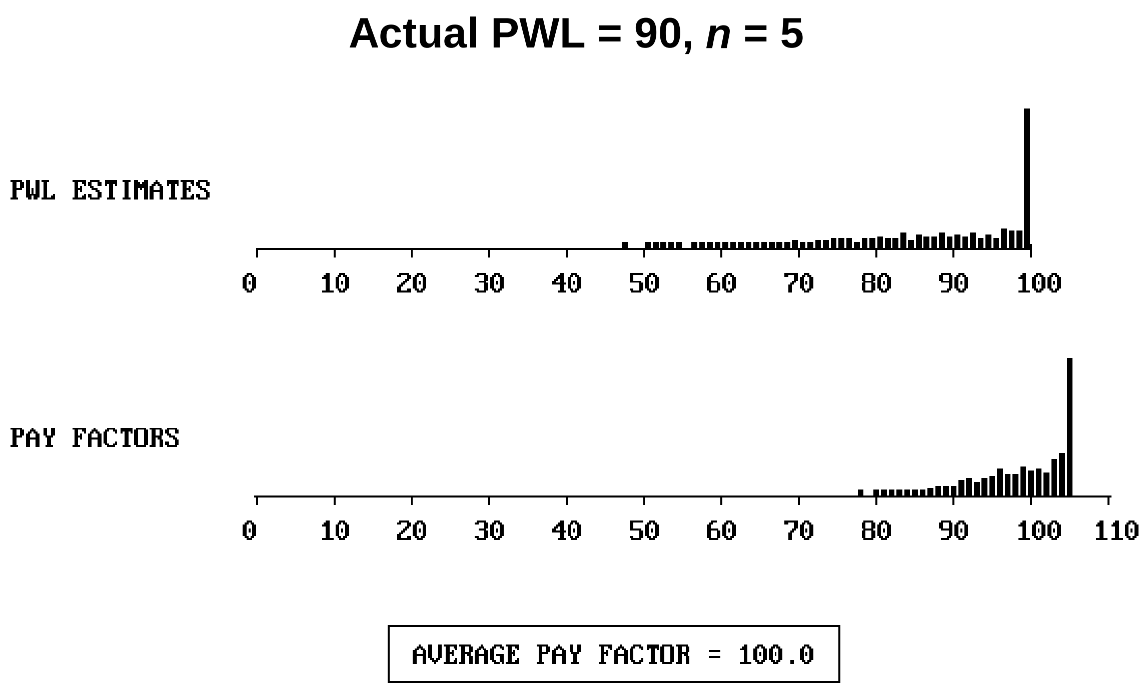


Figure 55c. Distribution of actual PWL = 90, sample size = 5, and the resulting payment factors.

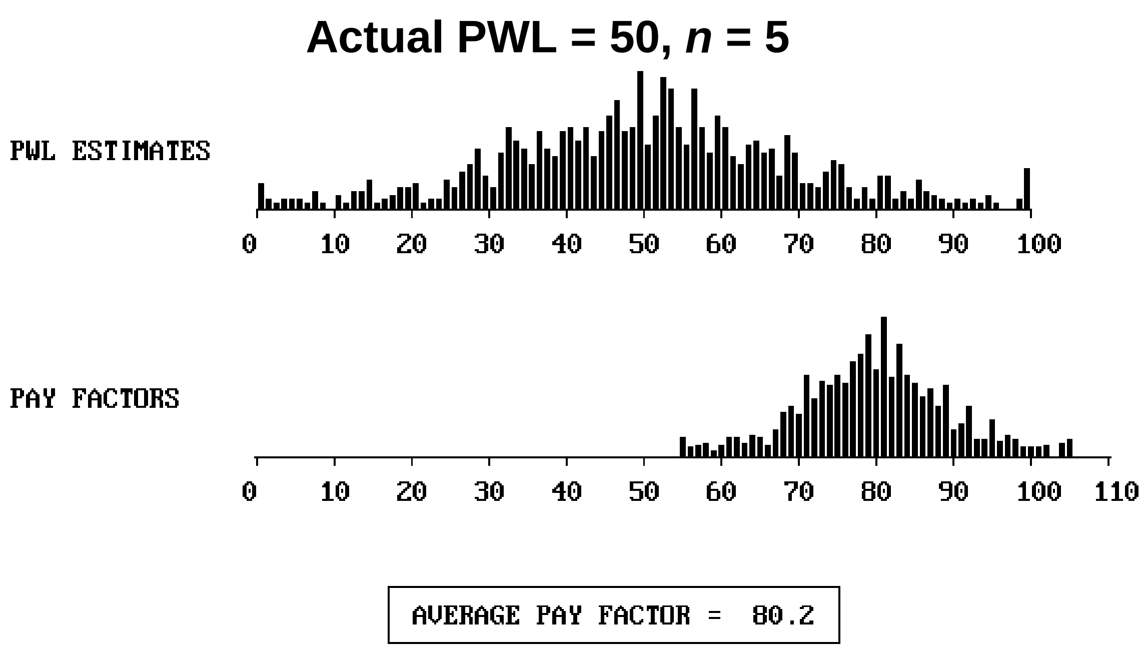


Figure 55d. Distribution of actual PWL = 50, sample size = 5, and the resulting payment factors.

Actual PWL = 90, *n* = 10

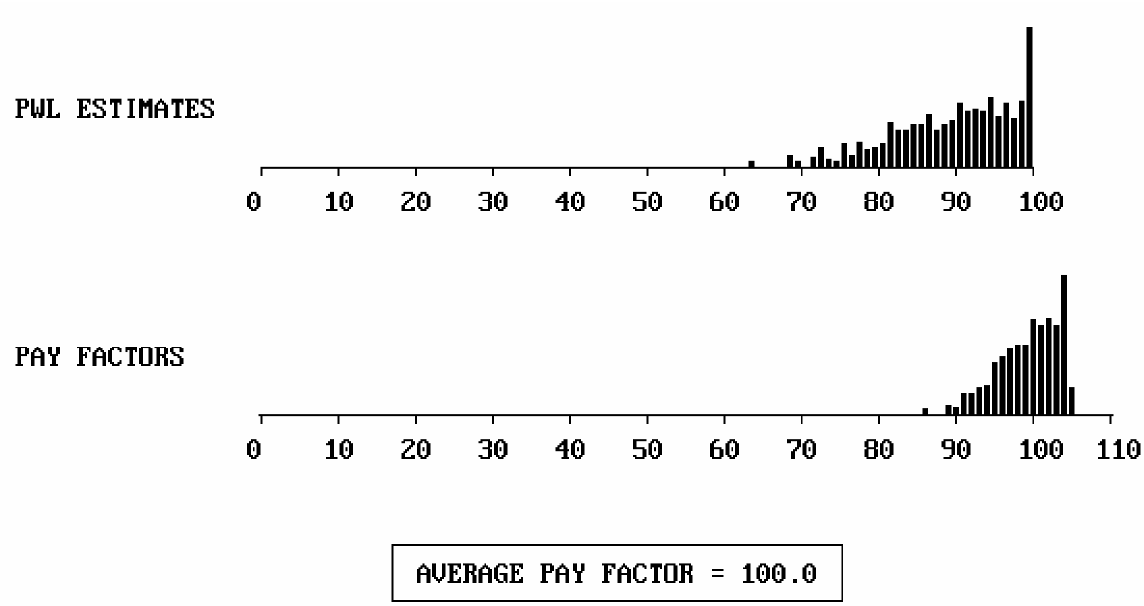


Figure 55e. Distribution of actual PWL = 90, sample size = 10, and the resulting payment factors.

Actual PWL = 50, *n* = 10

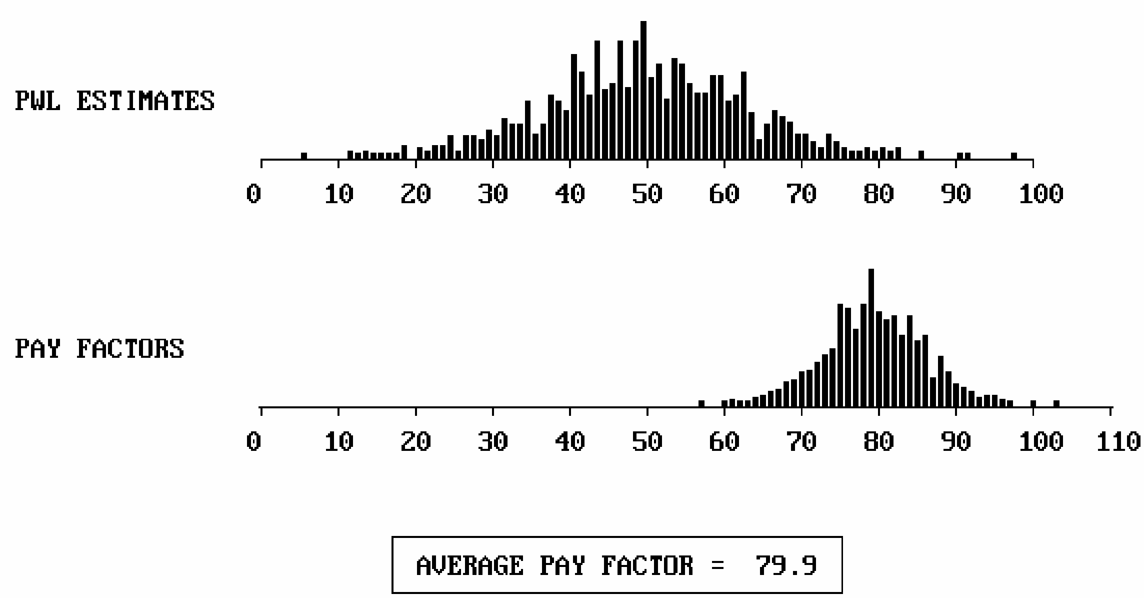


Figure 55f. Distribution of actual PWL = 50, sample size = 10, and the resulting payment factors.

AAD Payment Factors

Computer simulation was also used to generate EP curves for a payment equation that based payment on the estimated AAD value. Since AAD is an unbiased estimator of the actual AAD value, the EP curve should follow the payment equation exactly. For the purposes of the simulation, any payment equation could have been used.

Since it is known that the expected value for the distribution of absolute deviations for a normal distribution is 0.798σ, a logical equation for relating AAD to payment is:

$$\text{Pay} = 105 - 24.75(\text{AAD} - 0.798), \text{ and Pay} \leq 105\% \tag{19}$$

This equation provides a maximum payment of 105 percent for an AAD = 0.798σ (i.e., the value for a normal distribution centered on the target). This payment equation was used in the comparison analysis. It should be noted that this equation was not intended to provide exactly the same EP curve as the PWL payment equation.

Figure 56 shows an EP curve for a sample size = 5, using equation 19 to determine the payment factor. It is shown in the figure that the EP curve follows a straight line. This is true because the payment equation is a straight line and AAD is an unbiased estimator. Also, the same trends of increasing variability about the EP line with a smaller sample size and decreased variability with a larger sample size would hold true for AAD.

Table 38 shows the EP values from the plot in figure 56 and the payment factor that should have been received for a population with the AAD value shown. The final column shows the difference between the EP and the correct payment. For lower AAD populations (i.e., those with higher quality), the EPs are lower than called for in the payment equation (equation 19). As the AAD value increases, the EP values come in line with the correct values. This issue may be related to the selection of the payment equation rather than the AAD estimation process. This issue is discussed in the section on risks.

Matching PWL and AAD Payment Equations

Although at first PWL and AAD seem very different, they are both based on the assumption that the population being sampled is normal. All normal distributions can be related to the standard normal distribution (i.e., $\mu = 0.0$ and $\sigma = 1.0$) by transforming the measurement axis to standard deviation units by the relationship:

$$Z = \frac{X - \mu}{\sigma} \tag{20}$$

where: Z = number of standard deviation units from the population mean

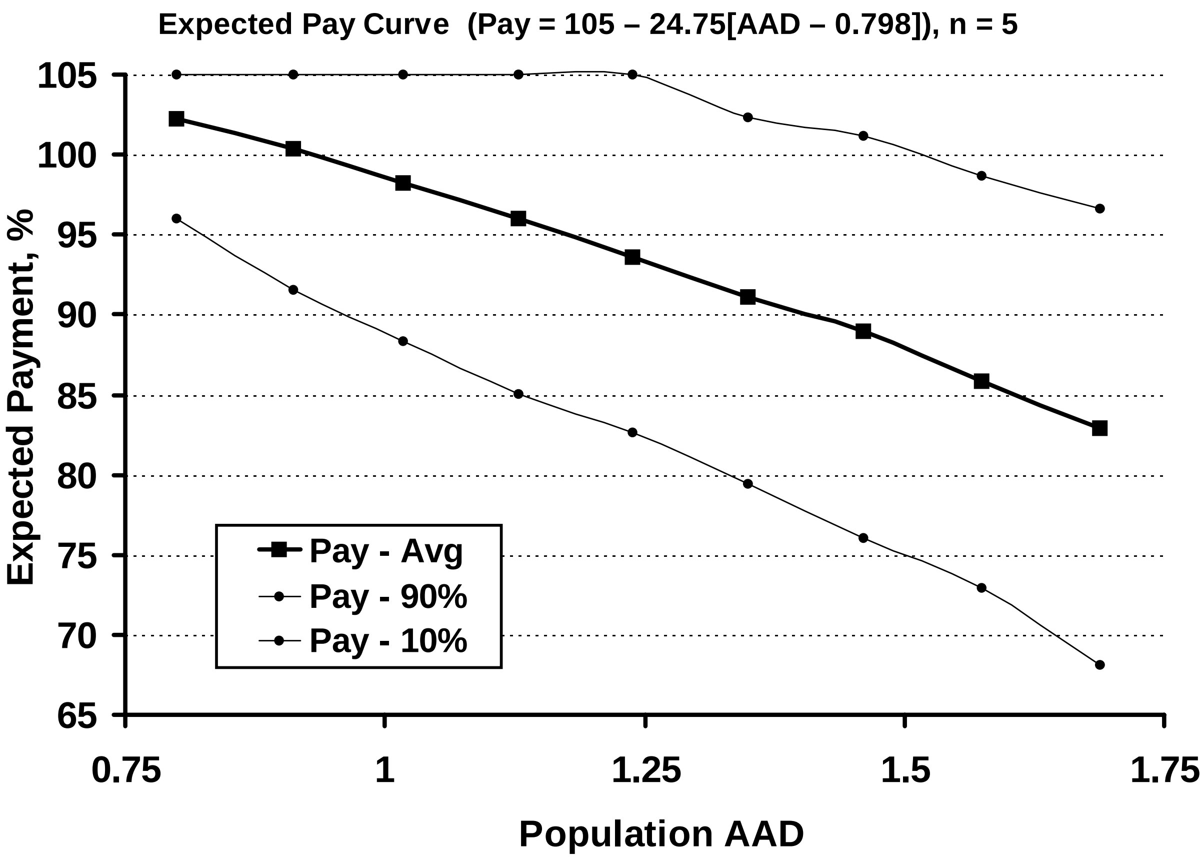
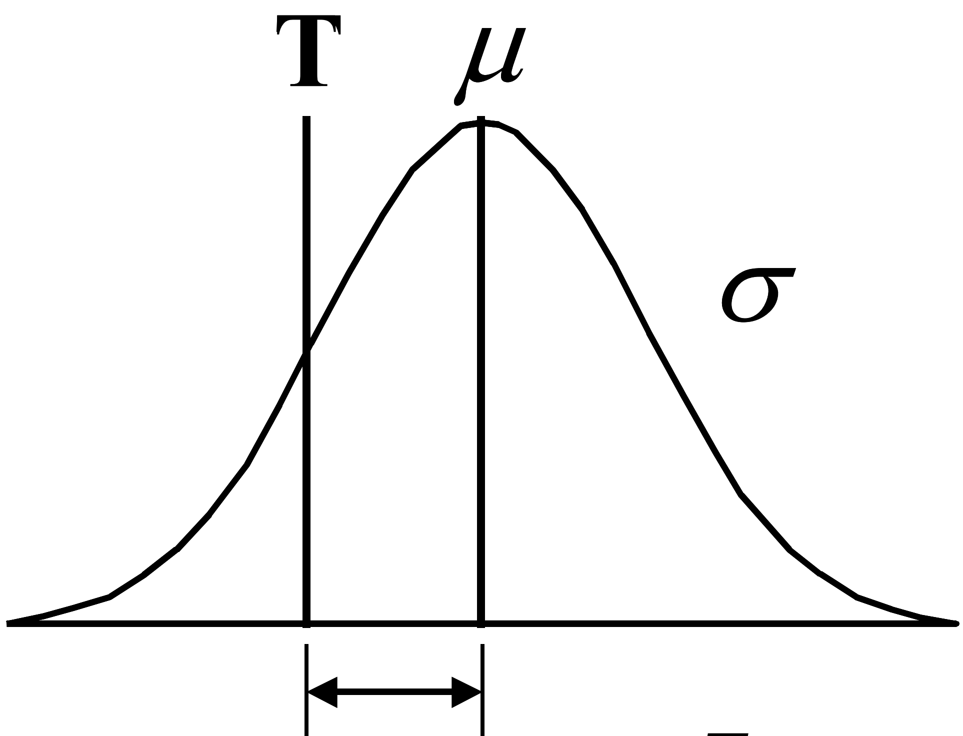


Figure 56. EP curve with the 90th and 10th payment percentiles for the AAD payment schedule.

Table 38. Simulated EP factors and correct payment factors based on AAD.

Actual AAD	Expected Payment	Correct Payment	Difference
0.798	102.24	105.00	–2.76
0.912	100.40	102.18	–1.78
1.018	98.21	99.55	–1.34
1.128	95.96	96.83	–0.87
1.238	93.60	94.11	–0.51
1.349	91.06	91.36	–0.30
1.461	88.93	88.59	+0.34
1.574	85.89	85.79	+0.10
1.688	82.90	82.97	–0.07

A similar approach can be taken for AAD. If AAD is measured in standard deviation units, then AAD can be related to the number of standard deviation units that the population mean is offset from the target value. This works for any normal distribution as long as AAD is measured in standard deviation units. In this way, AAD computer simulation results can be applied to any normal distribution in much the same manner that the standard normal curve can be used to calculate areas under any normal distribution. Figure 57 illustrates how this concept works. Z_{Targ} represents the distance that the population mean (μ) is offset from the target value (T) when measured in units of standard deviation (σ).



$$Z_{Targ} = \frac{\mu - T}{\sigma}$$

Figure 57. Illustration of measuring AAD in standard deviation units (Z_{Targ}) from the mean.

If we also measure the specification limits to be used in calculating PWL in terms of standard deviation units from the target value, then once the PWL specification limits are set, we can develop a relationship between any given PWL value and its equivalent AAD value for the same normal population.

In the simulation studies that were conducted to compare PWL results with AAD results for the same normal population, the PWL specification limits were set at $\pm 1.645\sigma$ from the target value, which corresponds to 90 PWL for a population centered on the target. Since it is now possible to determine AAD and PWL values for the same offset from the target (Z_{Targ}), an empirical relationship can be developed that for any normal population can convert any given PWL value to its corresponding AAD value, and vice versa. Equations 21 and 22 show this approximate empirical relationship in which all the terms have already been defined:

$$\mathrm{AAD} = Z_{\mathrm{Targ}} + 0.798e^{-1.613(Z_{\mathrm{Targ}})^{1.198}} \tag{21}$$

$$Z_{\mathrm{Targ}} = \mathrm{AAD} - 0.798e^{-3.12(\mathrm{AAD}-0.798)^{0.639}} \tag{22}$$

First, it should be noted that it is known that AAD = 0.7979σ if the population mean is centered at the target value. Therefore, since the AAD value for a population is approximately 0.798 times the population standard deviation, the only way that the AAD value could be zero is if the standard deviation was zero. Also, the smallest population AAD value is 0.798 times the population standard deviation. Therefore, if the population AAD value is measured in standard deviation units, then 0.798 is the smallest population AAD that is possible, thus, AAD ≥ 0.798 is a requirement.

Equations 21 and 22 allow the corresponding PWL value to be obtained for a given AAD value, and vice versa. For example, for a population with an AAD of 1.00, the offset of the population mean from the target value (Z_{Targ}), measured in standard deviation units, can be calculated as 0.7403 by using equation 22. Since the specification limits are assumed to be set at $\pm 1.645\sigma$, PWL can be determined from tables of areas under the standard normal curve using figure 58.

Tables 39 through 41 show the relationships between Z_{Targ}, measured in standard deviation units, and the AAD and PWL values for a normal population, provided that the PWL specification limits are two-sided and are set at $\pm 1.645\sigma$. The two-side specification limits are a requirement since AAD is always two-sided because it is measured as the deviation from a target value. With these tables, it is easy to identify the appropriate PWL and AAD values for a given normal population, provided that the specification limits are set at $\pm 1.645\sigma$.

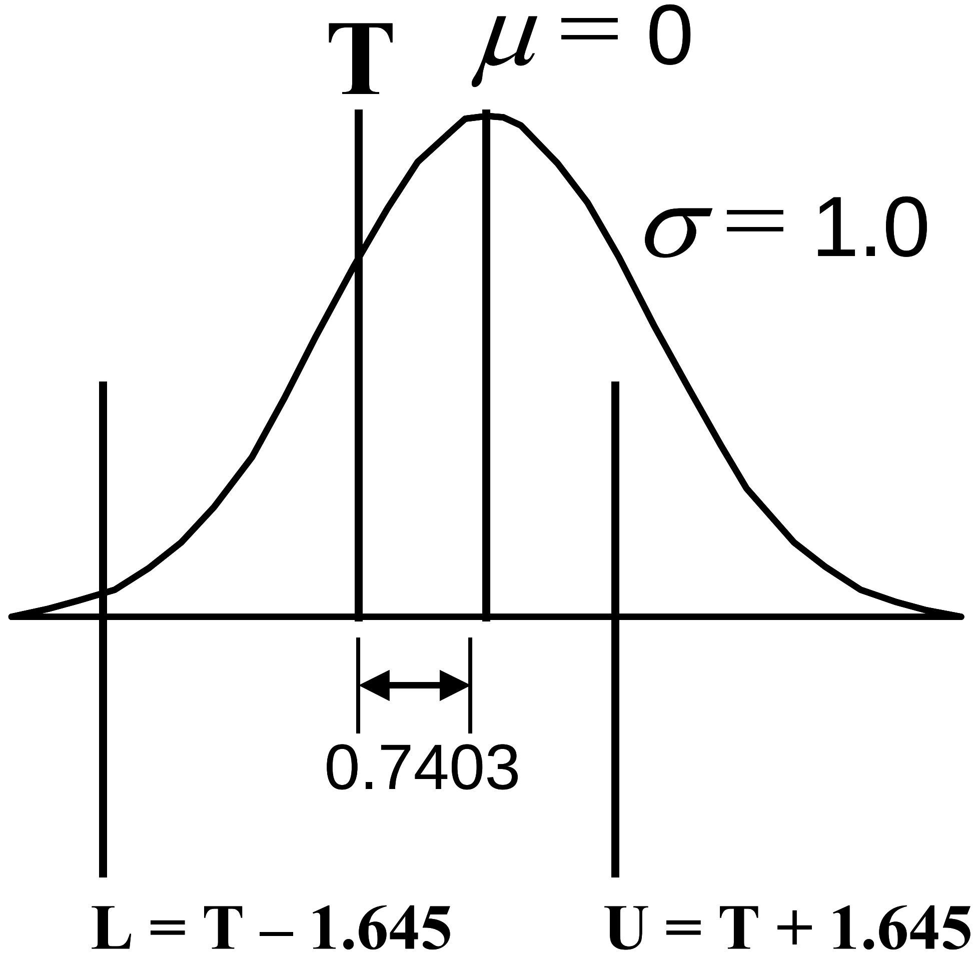


Figure 58. Example illustrating the PWL specification limits and the offset in σ units between the population mean and the target.

Table 39. Relationship between Z_{Targ}, AAD, and PWL for a normal population when the PWL specification limits are set at 1.645σ.

Z_{Targ}	PWL (Lower)	PWL (Upper)	PWL (Total)	AAD
0.00	95.00	95.00	90.00	0.798
0.10	95.95	93.88	89.83	0.820
0.20	96.75	92.58	89.33	0.831
0.30	97.41	91.07	88.48	0.845
0.40	97.96	89.34	87.30	0.866
0.50	98.40	87.39	85.79	0.895
0.60	98.76	85.20	83.96	0.933
0.70	99.05	82.77	81.82	0.979
0.80	99.28	80.09	79.37	1.032
0.90	99.45	77.19	76.64	1.093
1.00	99.59	74.05	73.65	1.159
1.10	99.70	70.71	70.41	1.231
1.20	99.78	67.18	66.96	1.307
1.30	99.84	63.50	63.33	1.388
1.40	99.88	59.68	59.56	1.471
1.50	99.92	55.76	55.68	1.558
1.60	99.94	51.79	51.74	1.647
1.70	99.96	47.81	47.77	1.738
1.80	99.97	43.84	43.81	1.831
1.90	99.98	39.94	39.92	1.925
2.00	99.99	36.13	36.12	2.020
2.10	99.99	32.46	32.45	2.116
2.20	99.99	28.94	28.94	2.213
2.30	100.00	25.62	25.62	2.310
2.40	100.00	22.51	22.51	2.408
2.50	100.00	19.63	19.63	2.506
2.60	100.00	16.98	16.98	2.605
2.70	100.00	14.57	14.57	2.704
2.80	100.00	12.40	12.40	2.803
2.90	100.00	10.47	10.47	2.902
3.00	100.00	8.77	8.77	3.002
3.10	100.00	7.28	7.28	3.102
3.20	100.00	6.00	6.00	3.201
3.30	100.00	4.90	4.90	3.301
3.40	100.00	3.96	3.96	3.401
3.50	100.00	3.18	3.18	3.501

Table 40. Relationship between AAD, Z_{Targ}, and PWL for a normal population when the PWL specification limits are set at 1.645σ.

AAD	Z_{Targ}	PWL (Lower)	PWL (Upper)	PWL (Total)
0.798	0.0000	95.00	95.00	90.00
0.800	0.0476	95.47	94.49	89.96
0.850	0.3521	97.71	90.20	87.91
0.900	0.5137	98.46	87.10	85.56
0.950	0.6371	98.88	84.33	83.20
1.000	0.7403	99.15	81.72	80.86
1.100	0.9131	99.47	76.79	76.26
1.200	1.0603	99.66	72.06	71.72
1.300	1.1929	99.77	67.44	67.21
1.400	1.3164	99.85	62.88	62.72
1.500	1.4337	99.90	58.37	58.26
1.600	1.5469	99.93	53.91	53.84
1.700	1.6570	99.95	49.52	49.47
1.800	1.7649	99.97	45.23	45.20
1.900	1.8711	99.98	41.05	41.03
2.000	1.9761	99.99	37.03	37.01
2.250	2.2348	99.99	27.77	27.76
2.500	2.4900	100.00	19.90	19.90
2.750	2.7433	100.00	13.60	13.60
3.000	2.9954	100.00	8.84	8.84
3.250	3.2468	100.00	5.46	5.46
3.500	3.4978	100.00	3.20	3.20
3.750	3.7484	100.00	1.77	1.77
4.000	3.9989	100.00	0.93	0.93
4.250	4.2492	100.00	0.46	0.46
4.500	4.4994	100.00	0.22	0.22

Table 41. Relationship between PWL, AAD, and Z_{Targ} for a normal population when the PWL specification limits are set at 1.645σ.

PWL (Total)	AAD	Z_{Targ}	PWL (Lower)	PWL (Upper)
90.00	0.7980	0.0000	95.00	95.00
85.00	0.9119	0.5456	98.58	86.42
80.00	1.0186	0.7753	99.22	80.78
75.00	1.1277	0.9558	99.54	75.46
70.00	1.2380	1.1121	99.71	70.30
65.00	1.3492	1.2546	99.81	65.19
60.00	1.4610	1.3886	99.88	60.12
55.00	1.5737	1.5175	99.92	55.07
50.00	1.6878	1.6437	99.95	50.05
45.00	1.8047	1.7699	99.97	45.03
40.00	1.9253	1.8978	99.98	40.02
35.00	2.0518	2.0301	99.99	35.01
30.00	2.1863	2.1693	99.99	30.00
25.00	2.3326	2.3194	100.00	25.00
20.00	2.4967	2.4867	100.00	20.00
15.00	2.6889	2.6816	100.00	15.00
10.00	2.9314	2.9264	100.00	10.00
5.00	3.2930	3.2900	100.00	5.00
1.00	3.9740	3.9728	100.00	1.00

Using the information from tables 39 through 41, a computer simulation program was developed that generated the EP and the standard deviation of the individual lot payment factors for any value of either PWL or AAD and the corresponding payment equation. The program output also identifies both the PWL and AAD values for the population. This program makes it simple to compare PWL and AAD payment plans provided that equivalent payment equations are developed for AAD and PWL. Tables 39 through 41 allow the development of such equivalent AAD and PWL payment schedules.

As an example, suppose an agency was considering the use of PWL with the payment factor determined by equation 18. However, the agency would also like to evaluate a payment equation based on AAD that would provide for a given normal population the same EP as the PWL equation. This is not difficult to do and, indeed, was done in this project to compare PWL and AAD payment procedures for the same population.

First, equation 18 was used to determine the payment factor for a number of actual PWL values ranging from 90 to 10. Table 41 was then used to identify the AAD value that corresponded to each of the PWL values. This then allowed the payment factors to be associated with their respective AAD values. The results of this process are shown in table 42.

Table 42. AAD values equivalent to the corresponding PWL payment factors.

Payment Factor	Corresponding PWL Value	Equivalent AAD Value
100	90	0.798
95	80	1.019
90	70	1.238
85	60	1.461
80	50	1.688
75	40	1.925
70	30	2.186
65	20	2.497
60	10	2.931

The AAD program that was developed allows for the use of a compound payment equation that is made up of two straight lines with different equations. To illustrate this and to help match the EP curve of the PWL payment equation, it was decide to use two equations that met at AAD = 1.925. The program allows for an equation in the form:

$$\text{Pay} = A - B(\text{AAD} - 0.798) \quad (23)$$

It is easy to solve for the *A* and *B* coefficients using the values in table 42. Doing so resulted in the following AAD payment plan to match the PWL payment equation:

$$\text{Pay} = 100 - 22.18(\text{AAD} - 0.798) \qquad \text{when AAD} \leq 1.925 \quad (24)$$

$$\text{Pay} = 91.80 - 14.91(\text{AAD} - 0.798) \qquad \text{when AAD} > 1.925 \quad (25)$$

One should note that equations 24and 25 each provide a payment factor of 75 when AAD = 1.925, thus providing the desired continuity in the two payment equations.

Computer simulation was used to compare the performance of a PWL payment plan based on equation 18 with that of an AAD payment plan based on equations 24 and 25. The results of these simulations are shown in table 43 and figures 59 and 60. The plans have very similar EPs, with the AAD plan being slightly higher in the mean offset range of about 1.5 to 2.5. However, table 39 shows that this represents PWL values of about 55 and below. The PWL standard deviations increase to a peak at a mean offset of about 1.6 to 1.7, and are higher than those for AAD for the middle portion of the plot. The PWL peak corresponds to a PWL of 50, which was shown previously to be the point of maximum variability for PWL estimates.

Table 43. Results of the simulation of matched PWL and AAD payment equations for sample size = 5.

Mean Offset	PWL Results			AAD Results		
	PWL	Expected Payment	Standard Deviation	AAD	Expected Payment	Standard Deviation
0.000	90	100.26	5.40	0.798	99.76	5.05
0.775	80	95.05	7.22	1.019	94.96	7.03
1.112	70	89.74	8.61	1.238	89.99	7.76
1.389	60	85.41	9.26	1.461	85.03	8.39
1.644	50	80.15	9.52	1.688	80.71	8.21
1.898	40	74.48	9.41	1.925	76.55	8.00
2.169	30	69.76	8.91	2.186	71.78	7.35
2.487	20	65.02	7.79	2.497	66.57	6.99
2.926	10	60.11	5.51	2.931	60.14	6.65

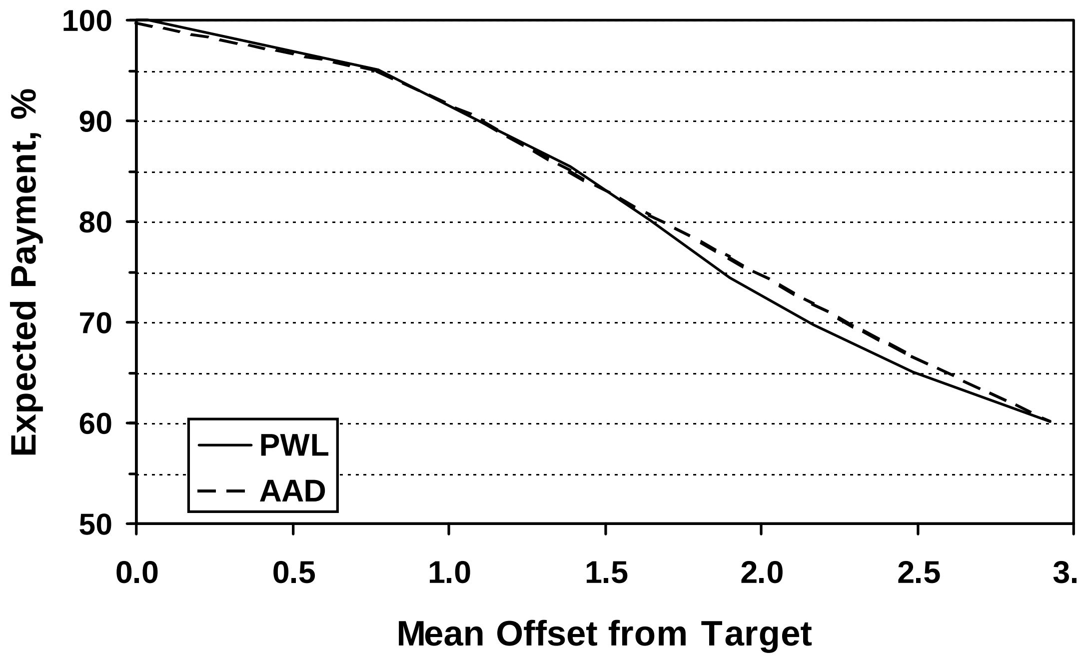


Figure 59. EP curves for matched PWL and AAD payment equations for sample size = 5.

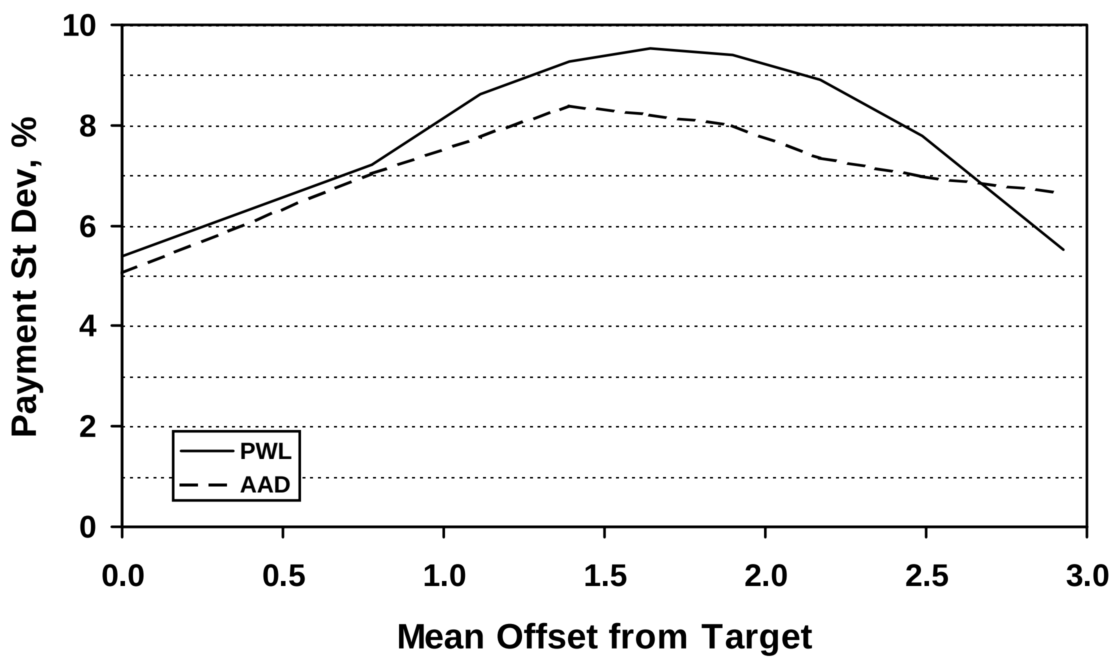


Figure 60. Standard deviations for individual payment factors for matched PWL and AAD payment equations for sample size = 5.

While there were some minor differences, the simulation study showed that when a single quality characteristic is used, it is possible to develop separate, but matched, PWL and AAD payment equations that provide similar EP curves and have reasonably close variability associated with the individual lot payment factors.

PAYMENT FOR MULTIPLE QUALITY CHARACTERISTICS

Since it is rare that only a single quality characteristic is used to determine the payment factor for a lot, it was also necessary to consider how the quality measures performed when multiple acceptance characteristics were used and to consider methods for combining these individual estimated quality measures into a combined, or composite, payment factor. While it was not within the scope of the original project, it was decided that, if multiple quality characteristics were to be investigated, it would also be necessary to consider potential correlations among these characteristics.

Simulating Correlated Variables

First it was necessary to develop computer routines for simulating correlated normal variables. These simulation routines were developed to simulate up to four correlated variables. The routines were then tested to verify that they were capable of generating samples with the intended level of correlation among the various variables.

Correlation is related to a pair of variables and the amount of correlation between the two variables is measured in terms of the correlation coefficient. Correlation coefficients can vary from +1.0 (as one variable increases, the other also increases with perfect correlation) to –1.0 (as one variable increases, the other decreases with perfect correlation). A correlation coefficient of 0.0 indicates that there is no correlation between the results of the two variables.

Figure 61 shows the output screens from the simulations of 5000 pairs of correlated normal variables, each with different selected correlation coefficients. The histograms on the left side of the plots show the two normal populations from which the correlated values were generated. The scatter plots on the right side of the figures show the data that were simulated. Figure 61(a) shows a situation where the two variables are not correlated. The resulting plot shows no pattern in the results when one variable is plotted against the corresponding second variable. In this simulation, the desired correlation coefficient was 0.000, while the correlation coefficient for the 5000 simulated pairs of values was –0.012, which is quite close to the desired value.

Figure 61(b) shows a similar plot for a desired correlation coefficient of +1.000. In this case, there is perfect correlation between the two variables. The increasing straight line at a 45-degree angle that is obtained when the two variables are plotted against one another shows this relationship. The correlation of the simulated value was also +1.000. A plot for a correlation coefficient of –1.000 would also be a straight line, but would decrease at a 45-degree angle. The remaining plots show comparisons of various levels of positive and negative correlation, and also show that the simulation routine does a good job of simulating the desired correlation.

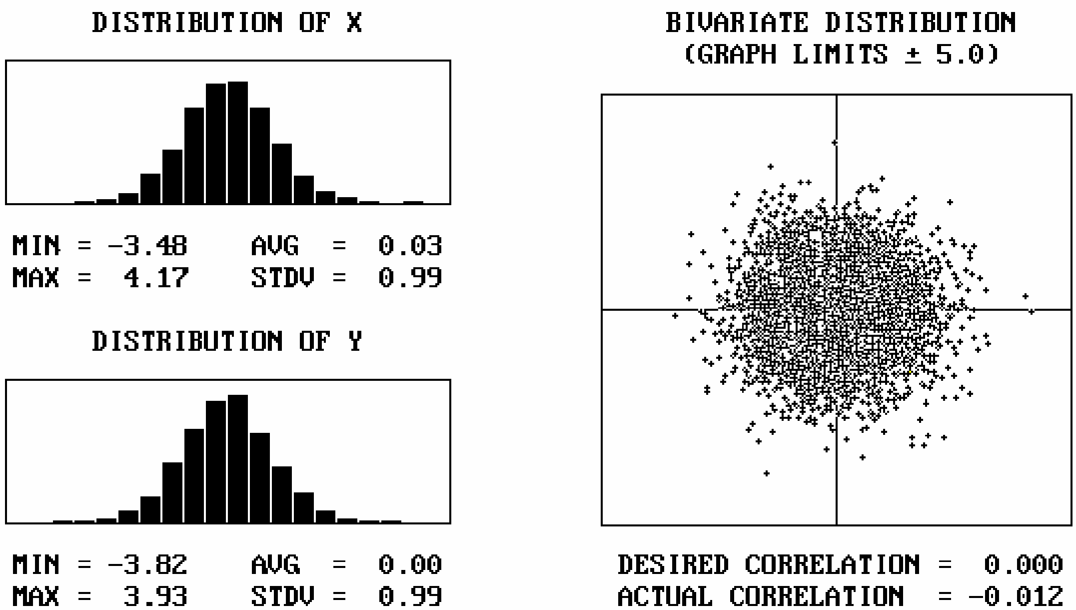


Figure 61a. Illustration 1 of two normal variables with various values for the correlation coefficient.

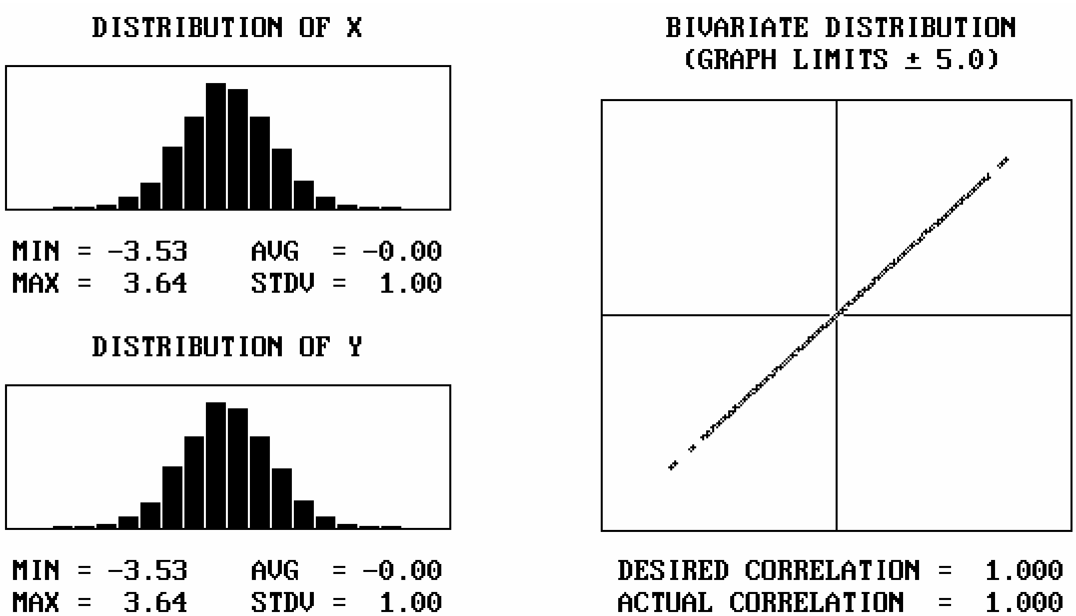


Figure 61b. Illustration 2 of two normal variables with various values for the correlation coefficient.

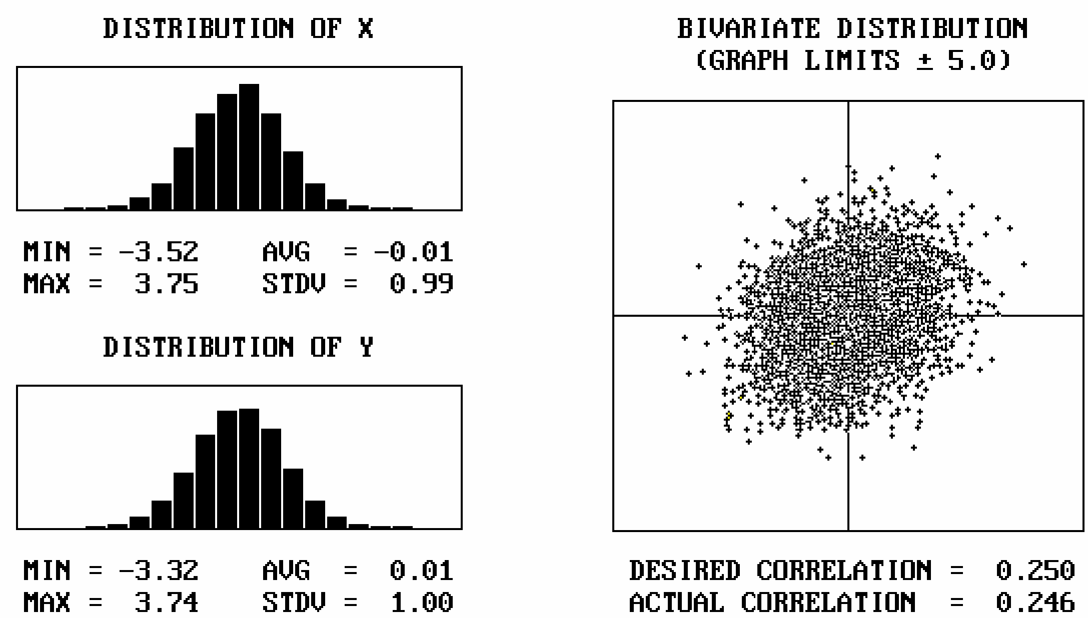


Figure 61c. Illustration 3 of two normal variables with various values for the correlation coefficient.

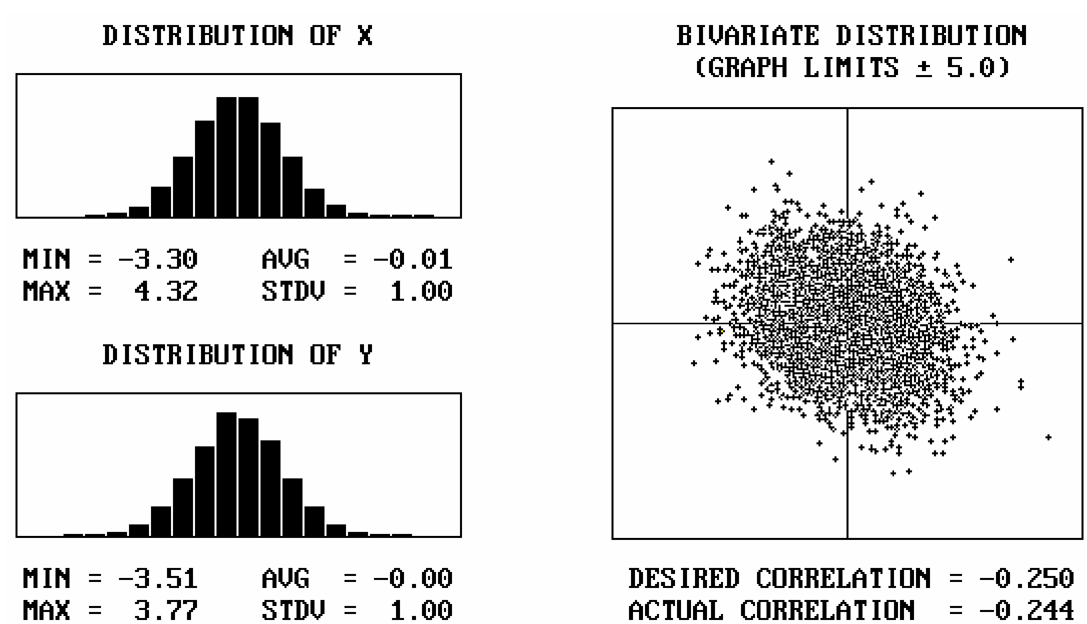


Figure 61d. Illustration 4 of two normal variables with various values for the correlation coefficient.

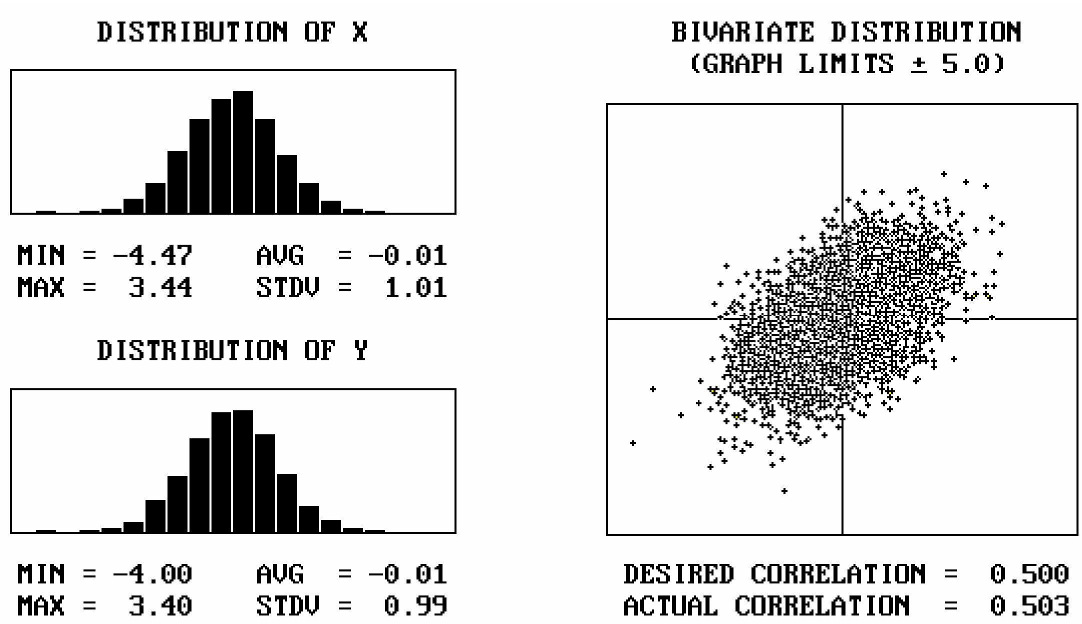


Figure 61e. Illustration 5 of two normal variables with various values for the correlation coefficient.

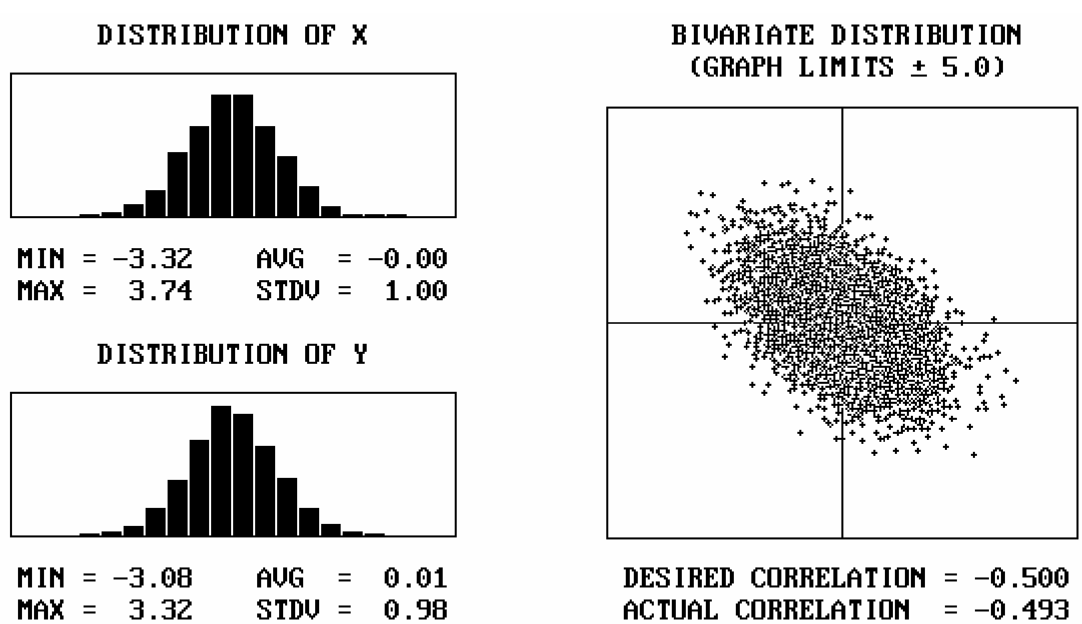


Figure 61f. Illustration 6 of two normal variables with various values for the correlation coefficient.

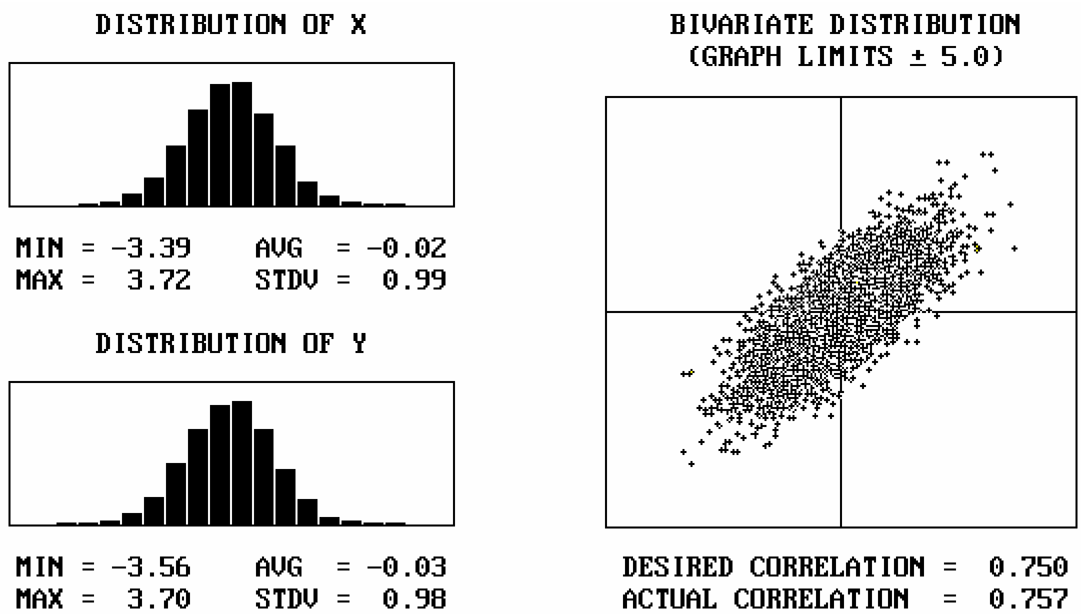


Figure 61g. Illustration 7 of two normal variables with various values for the correlation coefficient.

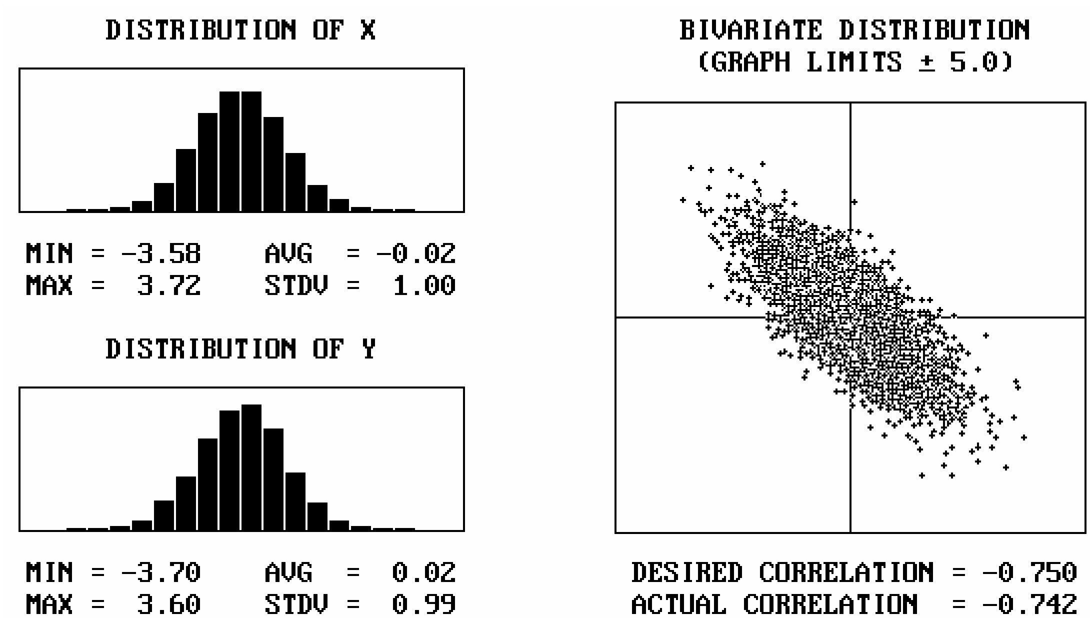


Figure 61h. Illustration 8 of two normal variables with various values for the correlation coefficient.

The simulation routine that was developed is capable of simulating up to four correlated normal variables. When more than two correlated variables are simulated, it is not possible to clearly show the correlation relationships as done in figure 61 for the two-variable case. However, tables 44 and 45 show the results of simulations for three and four correlated variables, respectively. While the results cannot easily be shown graphically, the tables show that the simulated data have correlation coefficients that compare closely with the desired values.

Table 44. Demonstration of the simulation of three correlated normal variables with selected values for correlation coefficients.

Correlated Variables	Desired Correlation Coefficient	Simulated Correlation Coefficient
Example 1		
Variable 1—Variable 2	+0.500	+0.497
Variable 1—Variable 3	+0.250	+0.259
Variable 2—Variable 3	+0.500	+0.510
Example 2		
Variable 1—Variable 2	+0.500	+0.500
Variable 1—Variable 3	–0.250	–0.262
Variable 2—Variable 3	–0.500	–0.509
Example 3		
Variable 1—Variable 2	–0.500	–0.500
Variable 1—Variable 3	–0.500	–0.496
Variable 2—Variable 3	–0.500	–0.504

Table 45. Demonstration of the simulation of four correlated normal variables with selected values for correlation coefficients.

Correlated Variables	Desired Correlation Coefficient	Simulated Correlation Coefficient
Example 1		
Variable 1—Variable 2	+0.500	+0.497
Variable 1—Variable 3	+0.250	+0.269
Variable 1—Variable 4	+0.500	+0.506
Variable 2—Variable 3	+0.500	+0.500
Variable 2—Variable 4	+0.250	+0.252
Variable 1—Variable 4	+0.500	+0.510
Example 2		
Variable 1—Variable 2	–0.250	–0.247
Variable 1—Variable 3	+0.500	+0.494
Variable 1—Variable 4	–0.250	–0.248
Variable 2—Variable 3	+0.500	+0.513
Variable 2—Variable 4	+0.250	+0.263
Variable 1—Variable 4	–0.500	–0.478

Correlated Quality Characteristics

Earlier in this chapter, the results from two simulation studies—one using PWL and equation 18 to determine payment and the other using AAD and equation 19 to determine payment—were compared. This comparison was expanded to address the case where two individual variables were considered when determining payment. While it is possible, and even likely, to have more than two variables involved in the acceptance and payment decision, it is not easy to picture the results when more than two variables are considered. Therefore, only the two-variable case was considered in this project. The results from two variables should be equally applicable to cases with three or four variables.

Two Populations of Equal Quality: The routines discussed above were used to compare the EP values and the amount of variability in the individual payment factors for both PWL and AAD when two correlated variables were involved. In the analyses for both PWL and AAD, the combined payment factor was taken as the weighted average of the two individual payment factors. This general payment relationship is shown in equation 26.

$$\text{Pay} = (W_1 \times PF_1) + (W_2 \times PF_2) \tag{26}$$

where: W_1 = weighting for variable 1
W_2 = weighting for variable 2
PF_1 = individual payment factor for variable 1
PF_2 = individual payment factor for variable 2

A sample size = 5 was used in the analyses. Three different correlation coefficients between the two variables were considered: 0.0, –0.5, and +0.5. In each analysis, the two individual variables had the same actual population: 70 PWL or 1.238 AAD. Various weightings (1.00/0.00, 0.90/0.10, 0.75/0.25, 0.50/0.50, 0.25/0.75, 0.10/0.90, and 0.00/100.00) were used to determine the combined payment factor.

Since, in each case, the two individual populations were of equal quality (i.e., 70 PWL or 1.238 AAD), the same EP should be obtained regardless of the weightings used. For the PWL case, the EP for each individual variable should be 55 + (0.5 × 70) = 90. For the AAD case, the EP for each individual variable should be 105 – [24.75 × (1.238 – 0.798)] = 94.1.

The results from the analyses are presented in table 46 and figures 62 and 63 for PWL, and in table 47 and figures 64 and 65 for AAD. The results for the 1.00/0.00 and 0.00/1.00 weightings are equivalent to the one quality characteristic using variable 1 or variable 2, respectively.

From table 46 and figure 62, it is shown that the EP values do not change as either the correlation coefficients or the variable weights change. The EP values are all around the value of 90.0 calculated from the payment equation. Therefore, PWL provides an unbiased estimate for the combined payment factor regardless of the weights selected or the amount of correlation between the variables.

However, the same is not true with respect to the variability of the individual simulated combined payment factors, as indicated by the standard deviation values in table 46 and figure 63. There is a noticeable difference among the standard deviation values for the three different correlation coefficients, with the difference being greatest at the 0.50/0.50 weighting of the two variables. Also, within the values for each correlation coefficient, there is a similar trend of decreasing variability as the 0.50/0.50 weighting is approached.

Although they are both related to the same mathematical relationship for the variance of the sum of two random variables, two different phenomena are at work in these two trends. It can be shown mathematically that if *a* and *b* are constants and *X* and *Y* are random variables, then:

$$\sigma_{aX+bY}^2 = a^2\sigma_X^2 + b^2\sigma_Y^2 + 2ab\sigma_{XY} \tag{27}$$

Table 46. Results of PWL simulation analyses for two correlated normal variables.

Weightings, Var_1/Var_2	Correlation Coefficient					
	0.00		+0.50		–0.50	
	Expected Pay	Std. Dev.	Expected Pay	Std. Dev.	Expected Pay	Std. Dev.
1.00/0.00	90.62	8.75	90.02	8.75	90.05	8.52
0.90/0.10	89.84	7.96	89.67	8.00	90.04	7.27
0.75/0.25	90.11	6.89	90.12	7.58	90.00	5.94
0.50/0.50	90.03	6.16	90.12	7.46	89.99	4.96
0.25/0.75	89.72	6.82	90.02	7.66	90.14	6.30
0.10/0.90	89.62	8.09	89.91	7.99	90.18	7.85
0.00/1.00	90.24	8.59	89.76	8.75	89.82	8.67

Table 47. Results of AAD simulation analyses for two correlated normal variables.

Weightings, Var_1/Var_2	Correlation Coefficient					
	0.00		+0.50		–0.50	
	Expected Pay	Std. Dev.	Expected Pay	Std. Dev.	Expected Pay	Std. Dev.
1.00/0.00	93.92	8.44	93.60	8.53	93.61	8.51
0.90/0.10	93.84	8.14	93.78	8.45	93.58	7.25
0.75/0.25	93.31	6.96	93.69	7.81	93.76	6.23
0.50/0.50	94.12	6.13	93.53	7.52	94.01	5.06
0.25/0.75	93.84	7.14	93.84	7.62	94.30	6.03
0.10/0.90	93.32	8.04	93.63	8.17	93.36	7.51
0.00/1.00	93.23	8.81	93.70	8.53	93.30	8.47

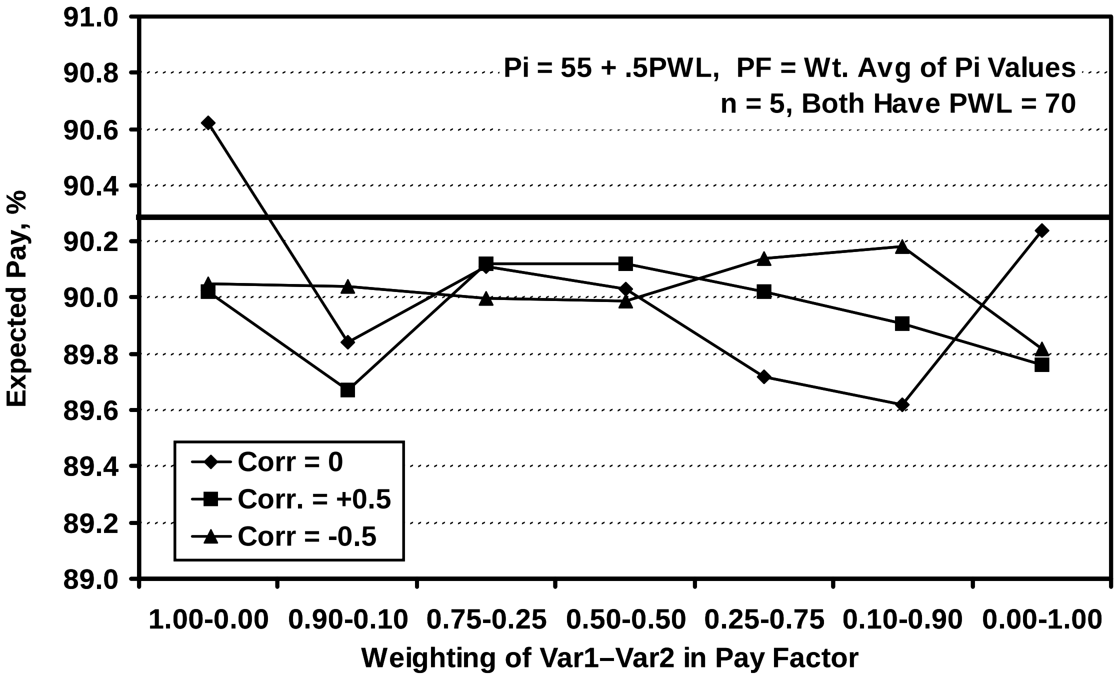


Figure 62. Expected combined weighted average payment factors for various weights and correlation coefficients based on PWL.

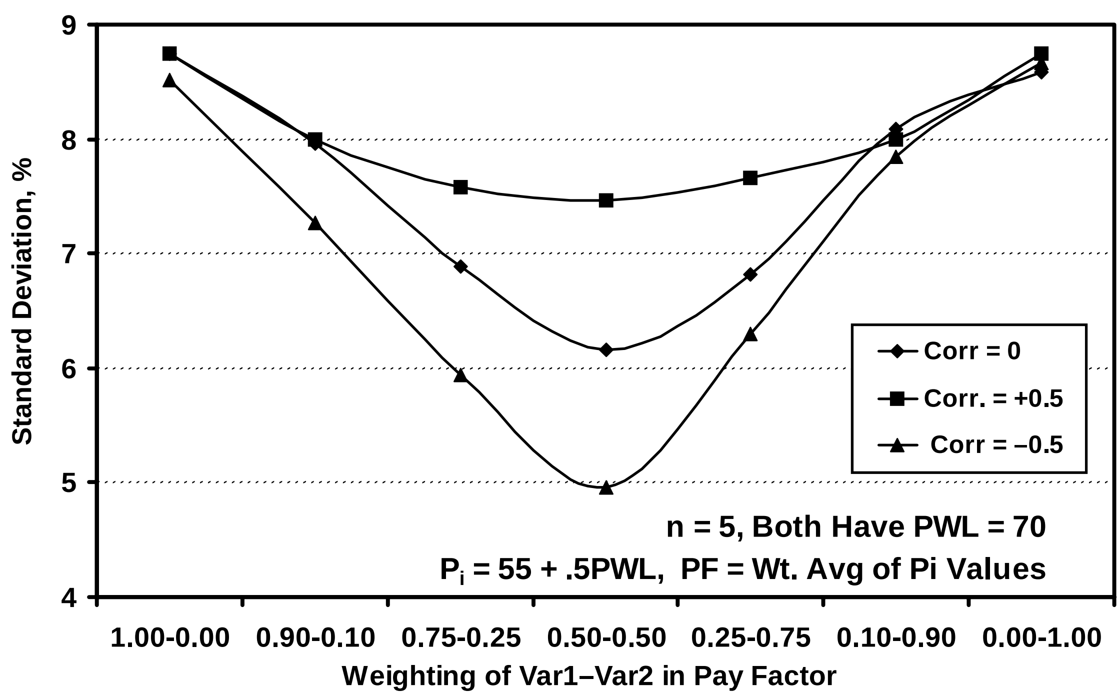


Figure 63. Standard deviations of weighted average payment factors for various weights and correlation coefficients based on PWL.

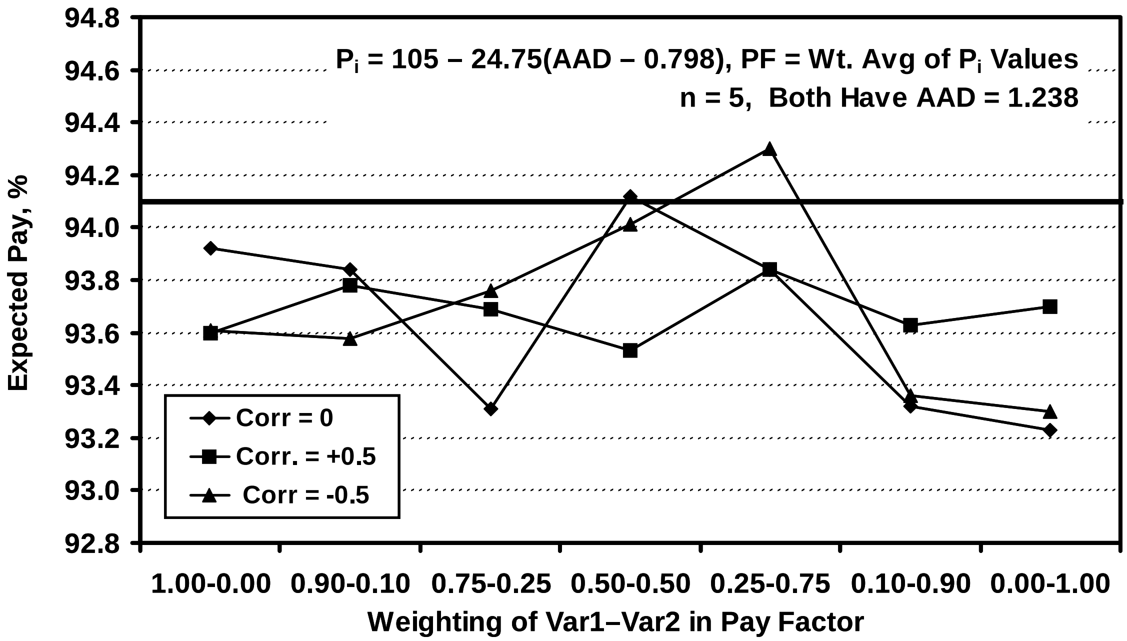


Figure 64. Expected combined weighted average payment factors for various weights and correlation coefficients based on AAD.

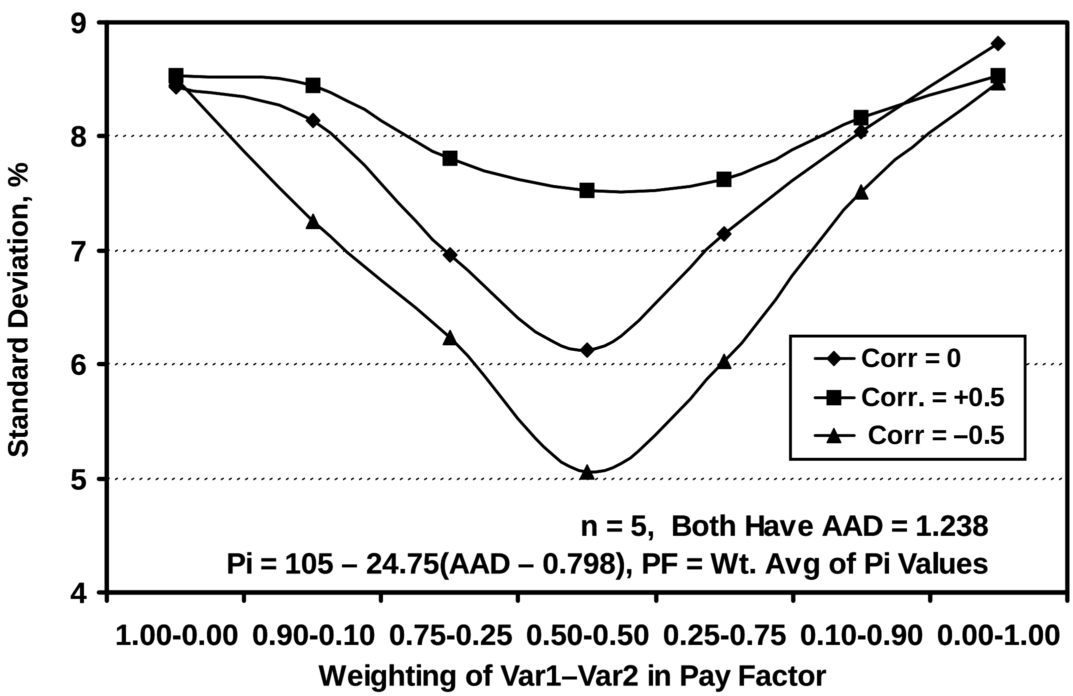


Figure 65. Standard deviations of weighted average payment factors for various weights and correlation coefficients based on AAD.

In equation 27, the last term is the covariance term and it is related to the correlation coefficient. The constants *a* and *b* can be related to the weightings (e.g., 0.50/0.50) that are multiplied by each of the random variables. Although equation 27 applies to the variability of the individual quality characteristic results, these results are used to determine PWL, which is linearly related to the payment factor by the payment equation. Therefore, equation 27 can be used to explain the standard deviation trends shown in figure 63.

In figure 63, for example, for a given set of weightings, such as 0.50/0.50, the standard deviation for +0.5 correlation is higher than for 0.0 correlation, which is higher than for –0.5 correlation. This follows directly from equation 27. For constant weightings, *a* and *b* in equation 27 remain the same for each correlation coefficient. Therefore, it is only the covariance term that changes the combined standard deviation value. For a correlation coefficient of 0.00, the covariance term in equation 27 also becomes 0.00. A positive correlation coefficient (and, therefore, a positive covariance term) results in a larger combined variability and, conversely, a negative correlation coefficient results in a smaller combined variability. Since the standard deviation also directly affects the AAD value, a similar standard deviation trend is shown for the AAD combined payment factors in table 47 and figure 65.

Equation 27 can also be used to explain the differences in standard deviation with respect to the weightings that are used. These weights are represented by the *a* and *b* constants in equation 27. For any given covariance value, the magnitudes of *a* and *b* will determine the differences in the combined standard deviations.

For example, take the case where the correlation coefficient is 0.00. In this event, the covariance term is eliminated from equation 27. Since the two variables have equal sample sizes, if measured in standard deviation units, they will also have equal variances for the purposes of calculating PWL and, therefore, payment. If we call this variance σ^2, then for the case of a 0.50/0.50 weighting, equation 27 leads to:

$$\sigma_{aX+bY} = \sqrt{0.5\sigma^2 + 0.5\sigma^2 + 0} = \sqrt{0.5\sigma^2} = 0.707\sigma \qquad (28)$$

For example, figure 63 shows that the standard deviation for PWL, with a correlation coefficient of 0.00 and weighting of 1.00/0.00 or 0.00/1.00, is about 8.7. Based on the calculations above, the standard deviation for a 0.50/0.50 weighting should be about $0.707\sigma = 0.707 \times 8.7 = 6.15$, which is essentially the same value shown in table 46. Similar calculations verify the shapes shown in figures 63 and 65.

Therefore, for two populations with equal PWL values, the combined payment factor, based on the weighted average of the two individual payment factors, has the same mean payment as for the individual variables, but with smaller variability. This is an unexpected result of using a weighted average approach for combining individual payment factors into one combined payment factor.

To consider populations with additional PWL values, simulations were run for the case of populations with 90 PWL, 70 PWL, and 50 PWL. The simulations were run with weightings of 0.50/0.50 and with correlation coefficients between the two variables ranging from –1 to +1. For both the PWL case (figure 66) and the AAD case (figure 67), the horizontal lines show that the EP value is not sensitive to the correlation between the individual variables, at least when the weighted average is used to determine the composite payment.

However, figures 68 and 69 indicate that the variability of the payment estimates is indeed related to the correlation between the two individual variables. This same trend is evident for both PWL and AAD. This means that while the average payment stays the same, the variability about the average payment increases as the correlation goes from –1.0 to + 1.0 for the two lower quality populations (i.e., 70 PWL and 50 PWL, and 1.238 AAD and 1.688 AAD), while it decreases slightly as the correlation approaches 0.0 for the higher quality population (90 PWL and 0.798 AAD). The relationship is different for the higher quality population since the upper boundaries of either 100 PWL or 0.0 AAD will have a greater impact than for the lesser quality populations.

Two Populations of Different Quality: All of the two-variable analyses up to now have been for the case where the two populations have the same quality level (i.e., the same PWL or AAD). Tables 47 through 51 show the results of simulation analyses for cases where the two individual populations have different PWL or AAD values and 50/50 weighting of the two variables when determining the combined payment factor. Tables 48 and 49 show that while, as expected, the quality levels of the individual variables affect the EP values, the EP values do not appear to be impacted by a moderate degree of correlation (i.e., –0.5, 0.0, and +0.5). Tables 50 and 51, however, show that the variability of the payment values may increase slightly with either positive or negative correlation between the two individual variables. This trend is evident in each case, except when one of the variables has the lowest quality simulated (i.e., 50 PWL or 1.688 AAD).

The values in tables 48 through 51 are plotted in figures 70 through 73. As shown previously, the payment biases using PWL are small and are centered at the correct value, regardless of the quality levels of the two populations. This is not the case for the payment biases using AAD, which are larger than for PWL and are negative for higher quality levels and positive for lower quality levels. For both PWL and AAD, the standard deviation values are lower for higher quality populations and there is little difference associated with the correlation coefficients. As the level of quality becomes lower, the standard deviations become larger and a difference develops with respect to the correlation coefficients. This trend is true for both PWL and AAD.

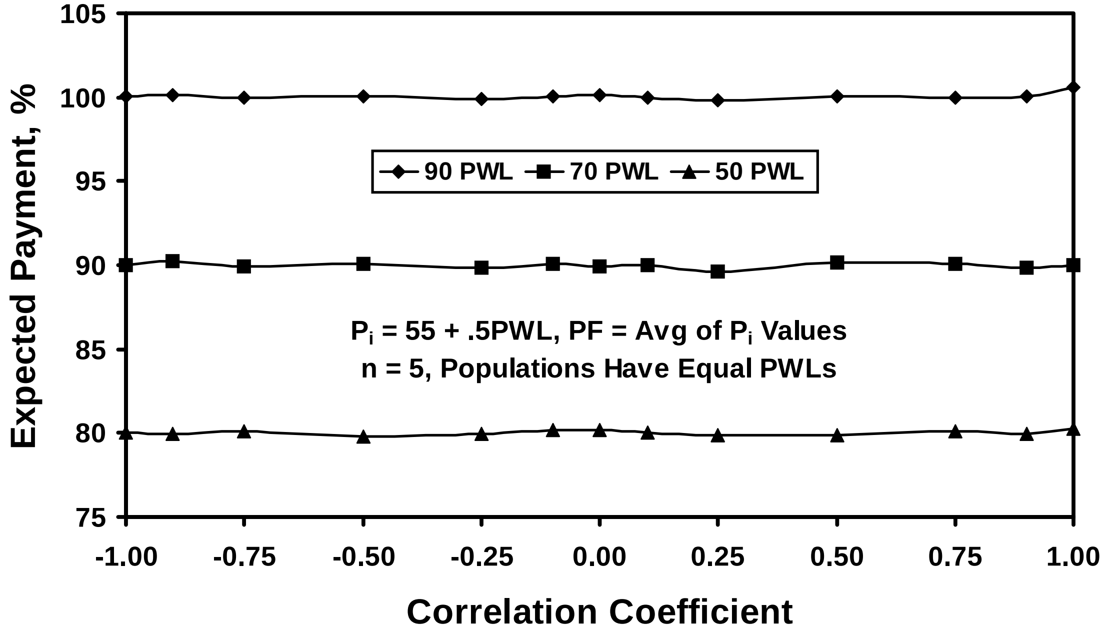


Figure 66. Expected average payment factors for two populations with various correlation coefficients based on PWL.

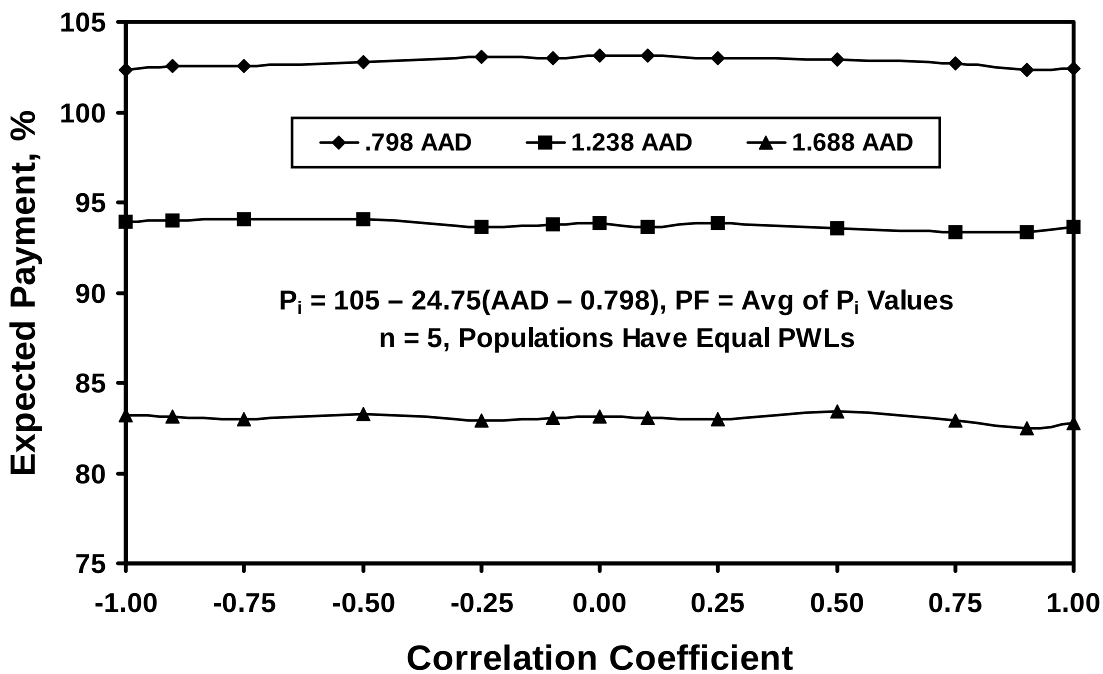


Figure 67. Expected average payment factors for two populations with various correlation coefficients based on AAD.

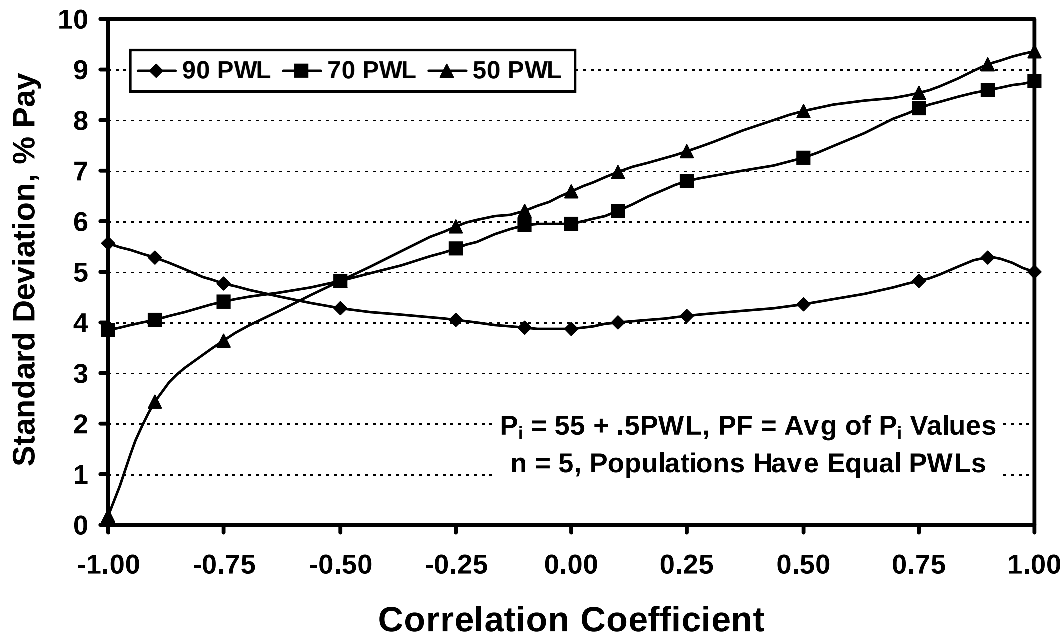


Figure 68. Standard deviations of individual payment factors for two populations with various correlation coefficients based on PWL.

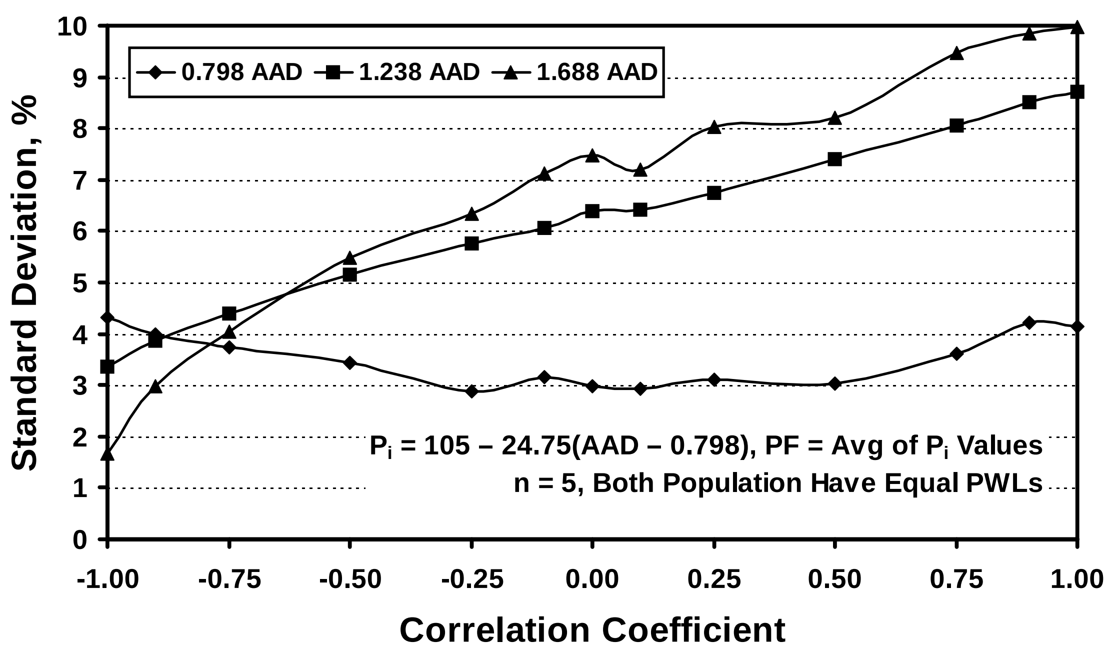


Figure 69. Standard deviations of individual payment factors for two populations with various correlation coefficients based on AAD.

Table 48. Bias in EP for two populations with equal PWL values and various correlation coefficients.

PWL–1	PWL–2	Correct Pay	Correlation Coefficient		
			–0.50	0.00	+0.50
90	80	97.50	–0.38	–0.30	+0.07
90	70	95.00	–0.06	+0.22	–0.07
90	60	92.50	+0.08	+0.16	+0.19
90	50	90.00	0.00	–0.17	+0.02
80	70	92.50	+0.21	+0.14	–0.08
80	60	90.00	+0.05	+0.11	–0.04
80	50	87.50	+0.13	–0.09	–0.13
70	60	87.50	+0.32	+0.14	–0.20
70	50	85.00	–0.05	–0.01	–0.04
60	50	82.50	–0.07	+0.16	–0.06

Table 49. Bias in EP for two populations with equal AAD values and various correlation coefficients.

AAD–1	AAD–2	Correct Pay	Correlation Coefficient		
			–0.50	0.00	+0.50
0.798	1.019	102.25	–1.11	–1.13	–1.35
0.798	1.238	99.55	–0.63	–0.54	–0.41
0.798	1.461	96.80	–0.33	–0.26	–0.01
0.798	1.688	94.00	–0.34	–0.34	–0.44
1.019	1.238	96.80	+0.96	+0.71	+0.92
1.019	1.461	94.05	+1.41	+1.13	+1.60
1.019	1.688	91.25	+1.51	+1.10	+1.47
1.238	1.461	91.35	+1.18	+1.24	+1.05
1.238	1.688	88.55	+1.53	+1.49	+0.74
1.461	1.688	85.80	+1.02	+1.32	+1.37

Table 50. Standard deviations of individual payment factors for two populations with equal PWL values and various correlation coefficients.

PWL–1	PWL–2	Correlation Coefficient		
		–0.50	0.00	+0.50
90	80	5.05	4.77	4.95
90	70	5.41	4.99	5.23
90	60	5.61	5.34	5.49
90	50	5.33	5.60	5.43
80	70	4.76	5.76	6.95
80	60	4.99	6.15	7.06
80	50	5.12	6.12	7.09
70	60	5.25	6.21	7.57
70	50	4.99	6.56	7.58
60	50	4.94	6.44	8.39

Table 51. Standard deviations of individual payment factors for two populations with equal AAD values and various correlation coefficients.

AAD–1	AAD–2	Correlation Coefficient		
		–0.50	0.00	+0.50
0.798	1.019	4.38	3.97	4.58
0.798	1.238	5.39	4.88	5.39
0.798	1.461	5.86	5.63	5.72
0.798	1.688	6.03	6.09	6.16
1.019	1.238	5.52	5.37	5.92
1.019	1.461	5.96	6.05	6.54
1.019	1.688	6.29	6.25	7.00
1.238	1.461	5.71	6.79	7.65
1.238	1.688	5.90	6.80	7.97
1.461	1.688	5.94	6.88	8.44

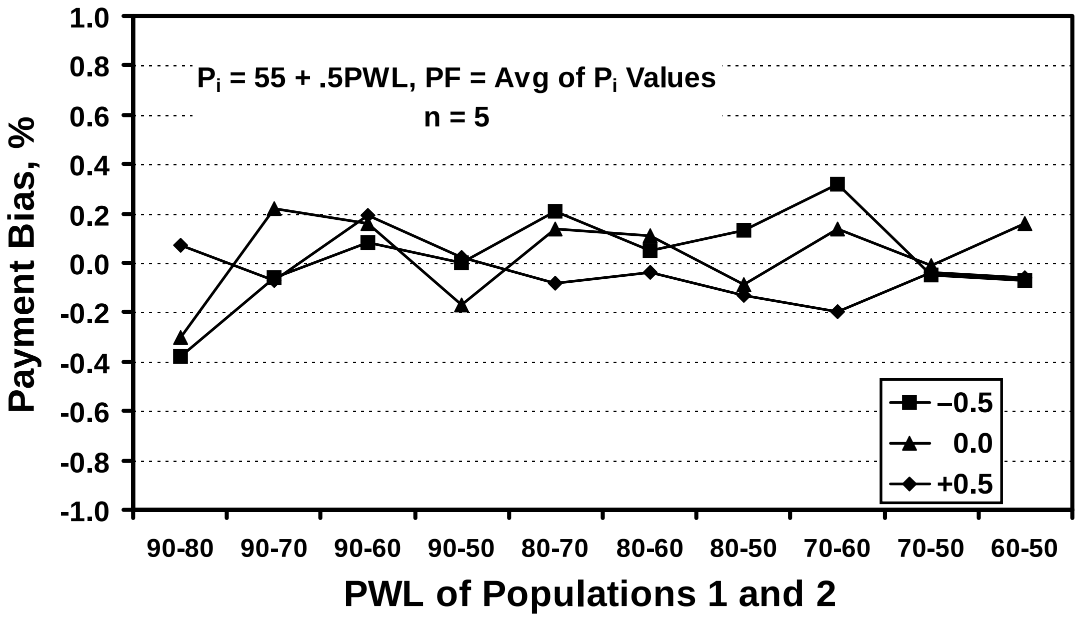


Figure 70. Bias for the average payment for two populations with various actual PWL values.

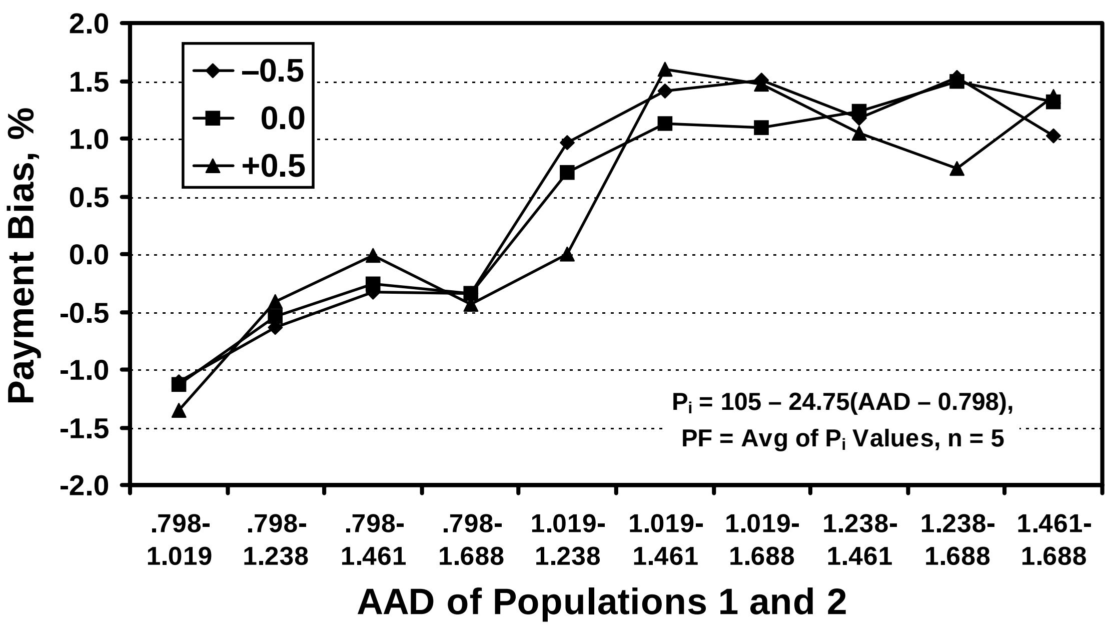


Figure 71. Bias for the average payment for two populations with various actual AAD values.

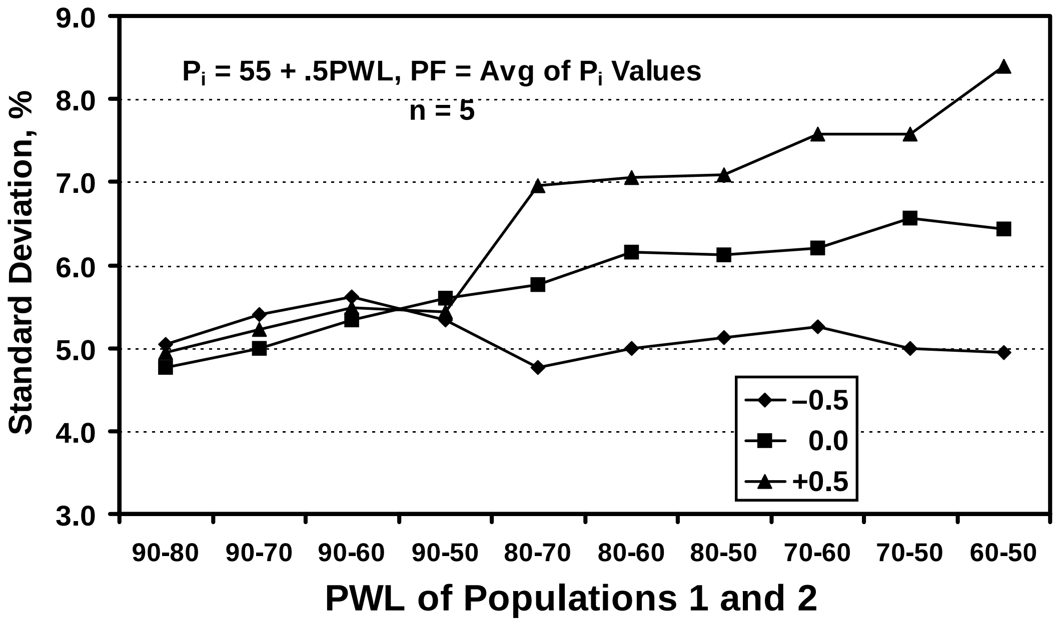


Figure 72. Standard deviation for the individual average payment values for two populations with various actual PWL values.

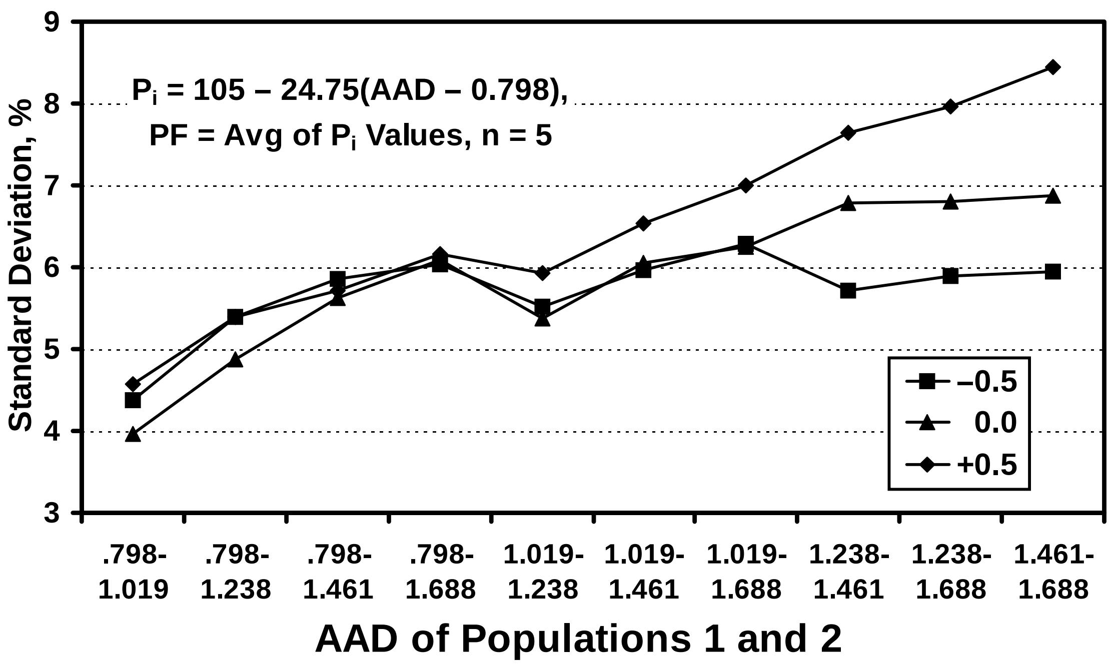


Figure 73. Standard deviation for the individual average payment values for two populations with various actual AAD values.

SELECTION OF THE QUALITY MEASURE

While the analyses indicated that, in general, the AAD quality measure performed about as well as the PWL quality measure, there were instances where the AAD measure had greater payment variability and exhibited some trends regarding payment bias. These issues, in themselves, would not be sufficient to eliminate AAD as a potential quality measure. However, there are some other drawbacks to AAD that make it a less appealing quality measure than PWL.

It is necessary to measure both center and spread when characterizing a population (lot) of material. Because of this, there are potential difficulties with the AAD quality measure. One drawback with AAD acceptance plans is that since lot variability is not really measured, a given lot AAD could come from a number of different populations. For example, the population could be centered at the target, but have a relatively large standard deviation (i.e., larger than the one that was assumed when developing the AAD acceptance limits or payment equation). Another population could have the same AAD by being centered off the target and having the same standard deviation that was assumed when developing the acceptance values. A third population could have the same AAD value by having a mean far from the target, but also by having a relatively small standard deviation. The fact that dissimilar populations can have the same AAD is shown by the distributions in figure 38 in chapter 6.

While some of these drawbacks may also apply to PWL (or PD) acceptance plans, such as the fact that a given PWL can represent many different populations, there are fewer drawbacks because both sample mean and standard deviation are measured in the PWL method. Also, since the PWL method can be used with both one-sided and two-sided acceptance properties, it is more versatile than the AAD method, which cannot be applied to one-sided limits.

Therefore, PWL (or its complement PD) was selected as the best quality measure for use in QA specifications. All remaining analyses were conducted only for the PWL (or PD) quality measure.

Now that the quality measure has been selected, its use needs to be further analyzed with respect to the risks associated with the small sample sizes that are typically used in highway materials and construction applications. Risks related to PWL and PD estimates, and their related payment factors, are discussed in chapter 9.

9. EVALUATING RISKS

INTRODUCTION

The concept of risk for acceptance plans, in general, and PWL, in particular, since that is our selected quality measure, is very similar to the concepts of accuracy and precision that are discussed in chapter 5 and the concepts of hypothesis testing and verification that are discussed in chapter 7.

Types of Risk

In hypothesis testing, there are two types of errors. A type I error occurs when a true null hypothesis is incorrectly rejected. A type II error occurs when an incorrect hypothesis is erroneously accepted. The probability of making a type I error is known as the α risk, while the probability of making a type II error is known as the β risk. When testing a null hypothesis, the α risk is also known as the level of significance.

The terms α and β have been applied in the highway construction industry for many years. However, these terms are related to a situation where the decision is either to accept or reject the hypothesis. In a highway construction scenario, the null hypothesis that is being tested is that the contractor's product meets the specification requirements. In such a scenario, the α and β risks apply as they do in any other hypothesis testing situation. Oftentimes, however, the decision is not simply to accept or reject the material, but to accept it at an incentive payment (positive price adjustment, or bonus) or at a disincentive payment (negative price adjustment, or penalty).

The Transportation Research Board (TRB) has published the following definitions for α and β risks:[(4)]

Seller's Risk (α) (also called a *type I error*): The probability that an acceptance plan will erroneously reject acceptable quality level (AQL) material or construction with respect to a single acceptance quality characteristic. It is the risk that the contractor or producer takes in having AQL material or construction rejected.

Buyer's Risk (β) (also called a *type II error*): The probability that an acceptance plan will erroneously fully accept (100 percent or greater) rejectable quality level (RQL) material or construction with respect to a single acceptance quality characteristic. It is the risk that the highway agency takes in having RQL material or construction fully accepted. (The probability of having RQL material or construction accepted (at any pay) may be considerably greater than the buyer's risk.)

When α and β risks are applied to materials or construction, they are really only appropriate for the case of acceptance or rejection decisions. When materials may not only be accepted or rejected, but may also be accepted at an adjusted payment, then the concept of α and β risks does not strictly apply. In such cases, which are perhaps more common than acceptance or rejection decisions, then much more involved analyses than just calculating α and β risks are necessary.

In an attempt to address the deficiencies in the use of α and β for price-adjustment acceptance plans, a new term (α_{100}) has been proposed. This is defined as the probability that AQL material will receive less than 100-percent payment. Just as with any term that tries to define a single point in a payment continuum, α_{100} provides only a little additional information regarding the realm of possible payments for a given quality of material.

OC and EP Curves

The only single term that truly attempts to consider the totality of possible payments for a given quality of material is EP. However, even this term is not sufficient since it considers only the average payment that a contractor can expect to receive for a very large number of production lots of a given quality level. EP fails to consider the amount of variability in the individual payment values that comprise the calculation of EP. To fully evaluate the risks in a price-adjustment acceptance plan, it is necessary to determine OC curves and EP curves, and also to investigate the amount of variability in the individual payment values about the EP curve.

TRB defines OC and EP curves as:[4]

OC Curve: A graphic representation of an acceptance plan that shows the relationship between the actual quality of a lot and either: (1) the probability of its acceptance (for accept/reject acceptance plans) or (2) the probability of its acceptance at various payment levels (for acceptance plans that include pay-adjustment provisions).

EP Curve: A graphic representation of an acceptance plan that shows the relationship between the actual quality of a lot and its EP (i.e., mathematical pay expectation, or the average pay the contractor can expect to receive over the long run for submitted lots of a given quality). (Both OC and EP curves should be used to evaluate how well an acceptance plan is theoretically expected to work.)

The definition of OC curves indicates that multiple curves might be plotted for the probability of receiving various levels of payment for a given actual quality level. The obvious difficulty with this is that if, as is recommended, an equation is used to calculate the payment factor, then there are an infinite number of OC curves that could be developed. Certain values are obvious candidates (e.g., the probability of receiving full payment or, if there is a remove-and-replace provision for very poor quality, the probability of requiring removal and replacement). It is difficult to view a plot with numerous OC curves (one for each of many selected payment levels) and to interpret it with a great deal of significance.

The EP curve, in essence, converts the family of OC curves into a single curve that represents the average payment that would be received in the long run for a given level of quality. This is the single curve that provides the most meaningful and useful information for a contractor or highway agency. However, even an EP curve is not sufficient since it does not consider the amount of variability in the individual payment values. This is why most of the analyses presented in this report consider not only bias or accuracy, but also standard deviation, to indicate the variability in addition to the average alone.

CALCULATING RISKS FOR A SINGLE QUALITY CHARACTERISTIC

Risks are traditionally determined separately for each individual quality characteristic. This process can be quite complicated, particularly if there are unusual provisions incorporated into the specification. Although not identified specifically as *risk analysis* at the time, much of the information presented in chapter 5 is directly related to the concept of risk. In this section, the material from chapter 5 that is related to the PWL quality measure is revisited and discussed in light of its application to risk analysis.

The computer simulation routines that were used to determine figures 9 through 12 were developed to calculate the information necessary to plot OC and EP curves when using either PWL or PD as the quality measure. The routines were actually developed for PD since the calculations are slightly less involved. While the programs can print the results for either PD or PWL, the calculations are performed for PD and are then converted to PWL if necessary.

Example

To illustrate the process of evaluating risks, the following payment equation, which is from the *AASHTO Quality Assurance Guide Specification* and which has been used many times in other chapters, is used:[3]

$$\text{Pay} = 55 + 0.5\text{PWL} \tag{29}$$

While this equation is not necessarily recommended, it has been used by a number of agencies. This equation is simple because it is a single straight line. It would be preferable to use either more than one straight line over different regions or to use some form of curved payment equation. This would allow for a shallower slope (i.e., lower payment reduction) for quality levels near the AQL and for a steeper slope that allows the price reductions to increase by a greater amount as the quality departs more from the AQL. For the purposes of illustration, equation 29 will work well. We also need to select values for AQL and RQL in terms of PWL. For this example, we will use 90 PWL as the AQL value and 60 PWL as the RQL value.

Also, suppose that the acceptance plan calls for removal and replacement of the material if the calculated PWL estimate is less than 60 (i.e., less than the RQL value). Therefore, the α risk (i.e., the risk that the AQL material will be rejected) is the probability that an AQL population will have an estimated PWL value of less than 60. This can be determined by simulating a large number of lots from an AQL population and determining the percentage of them that have PWL estimated values of less than 60.

The results of such a computer simulation of 1000 lots for a sample size = 4 are plotted in figure 74. The horizontal line on the plot in figure 74 indicates the α risk. It is the probability that AQL material will be rejected and is shown on the plot as the difference between 100-percent probability of acceptance and the probability of acceptance at AQL = 90 PWL. The α risk is about 0.025 (or 2.5 percent). This seems like a low risk, but remember that this is the risk of rejection. The curve in figure 74 is the probability of receiving at least some payment. While rejection is unlikely, there may be a significant chance of receiving less than full payment.

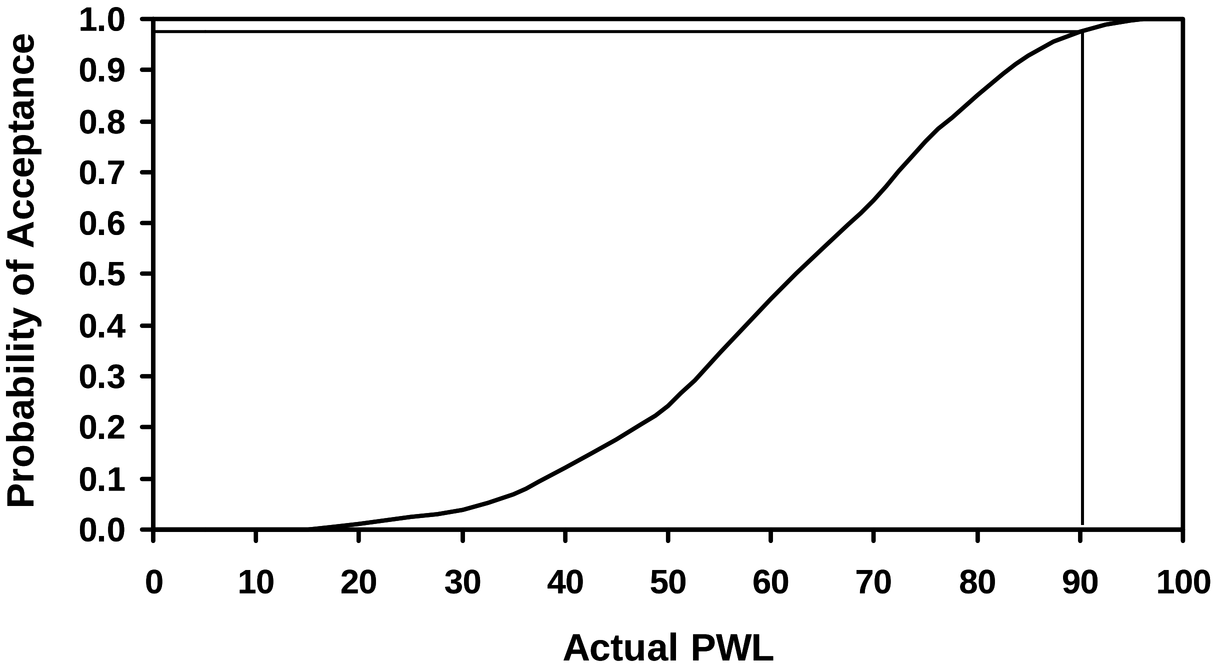


Figure 74. OC curve for an acceptance plan that calls for rejection if the estimated PWL is less than 60, sample size = 4.

In fact, the probability of receiving at least 100-percent payment can be determined in the same way as the probability of acceptance, except that the probability in question is the probability of having an estimated PWL $\geq$ 90. This value of 90 is calculated as the PWL value that yields 100-percent payment from equation 29.

The probability of receiving at least 100-percent payment is shown in figure 75, along with the probability of acceptance at any price. While the probability of receiving at least some payment is quite high for AQL material, the probability of receiving 100-percent or greater payment is only about 0.60 (or 60 percent). Therefore, for this example, the value of α_{100} would be about 1.00 – 0.60 = 0.40 (or 40 percent). That is the risk that the AQL material would receive less than full payment.

From this example, it appears that α is quite low at about 2.5 percent and that α_{100} is quite high at about 40 percent. However, neither of these numbers tells the full story. For one thing, they apply only to one specific level of quality—AQL = 90 PWL. Secondly, they address only two specific points on the payment continuum—receiving greater than zero payment and greater than or equal to 100-percent payment.

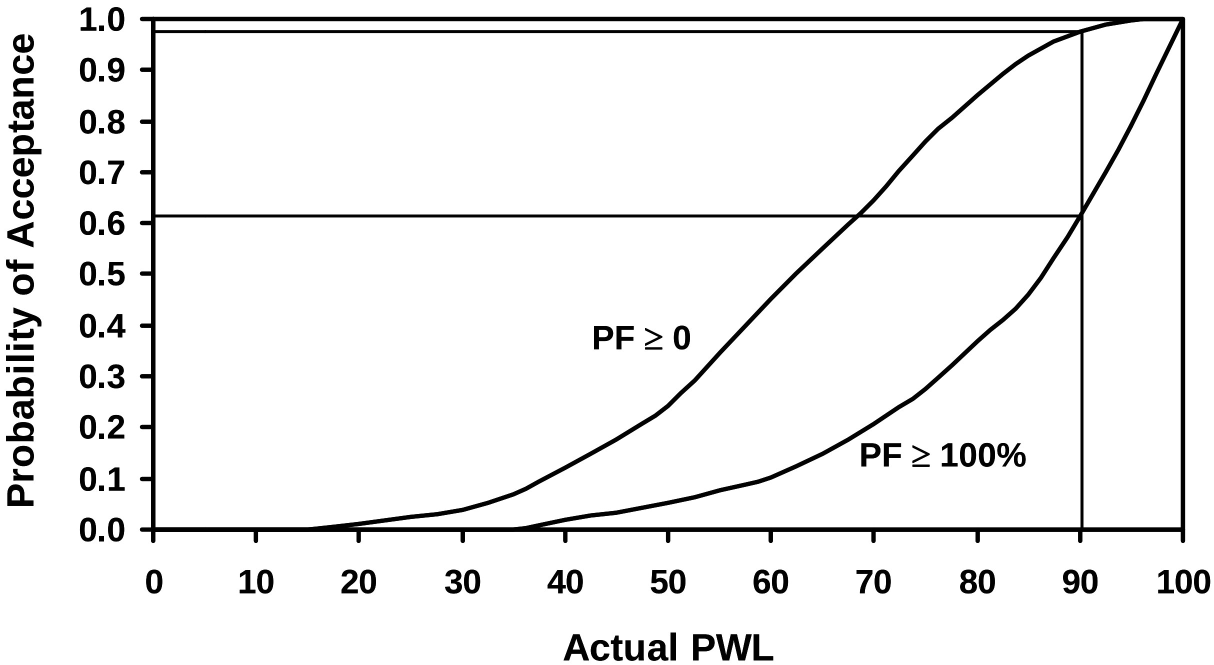


Figure 75. OC curves for the probabilities of receiving at least some payment and at least 100-percent payment, sample size = 4.

Neither of these pieces of information is sufficient to fully evaluate the risk associated with this payment plan. Many other curves, such as the probability of receiving at least 90-percent payment or of receiving at least 104-percent payment, could be added to the plot. In reality, there is a continuous payment region that runs from the curve for the probability of receiving at least some payment to the curve for the probability of receiving some high payment. The maximum payment of 105 percent can only be averaged if the population has 100 PWL; it cannot be plotted except as a point at the upper right-hand corner. However, the curve representing, say, 104.5-percent payment can be calculated and is plotted in figure 76, along with the curves for greater than zero payment and greater than or equal to 100-percent payment.

Figure 76 shows that the probability of receiving at least 104.5-percent payment for AQL material is approximate 47 percent. To at least some extent, this should help to balance the fact that there is about a 40-percent probability of receiving less than 100-percent payment. To evaluate the overall long-term performance of the acceptance plan and the potential payment risks for the contractor, it is necessary to determine the EP curve for the plan.

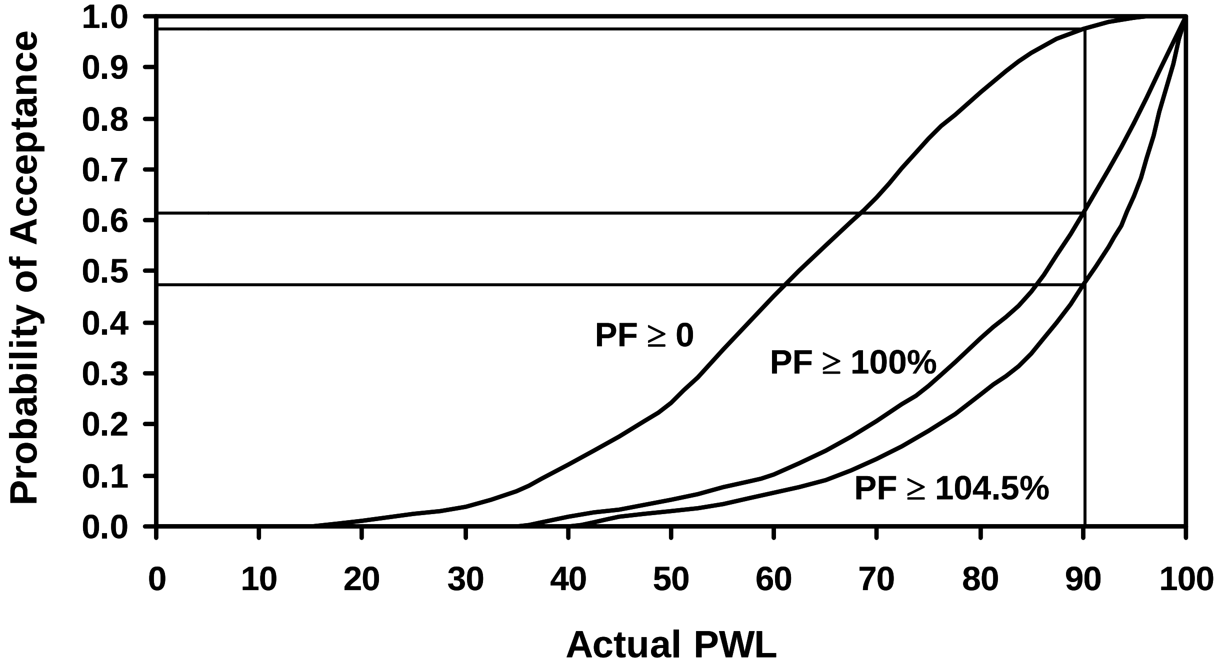


Figure 76. OC curves for the probabilities of receiving various payments, sample size = 4.

Figure 77 shows the EP curve for the payment schedule in equation 29, with the added provision that estimated PWL values of less than 60 receive no payment (i.e., they require removal and replacement at the contractor's expense). Figure 77 shows that the EP in the long run for AQL material is less than 100 percent. It is generally accepted that the EP for AQL material should be 100 percent. Therefore, the difference between the EP at the AQL in figure 77 and 100 percent can be thought of as a long-term payment risk for the contractor.

Since it has been shown that the sample PWL is an unbiased estimator of the true PWL of the population, the EP curve should exactly follow the shape of the payment equation. Indeed, that is what happened with the simulations that were done using this same payment equation in chapter 5. Why, then, does the EP at the AQL not equal the value of 100 that is given by the equation? When the EP does not equal the value stated by the payment equation, it usually means that some payment barrier has interfered with the ability of the EP to average out to the value in the payment equation.

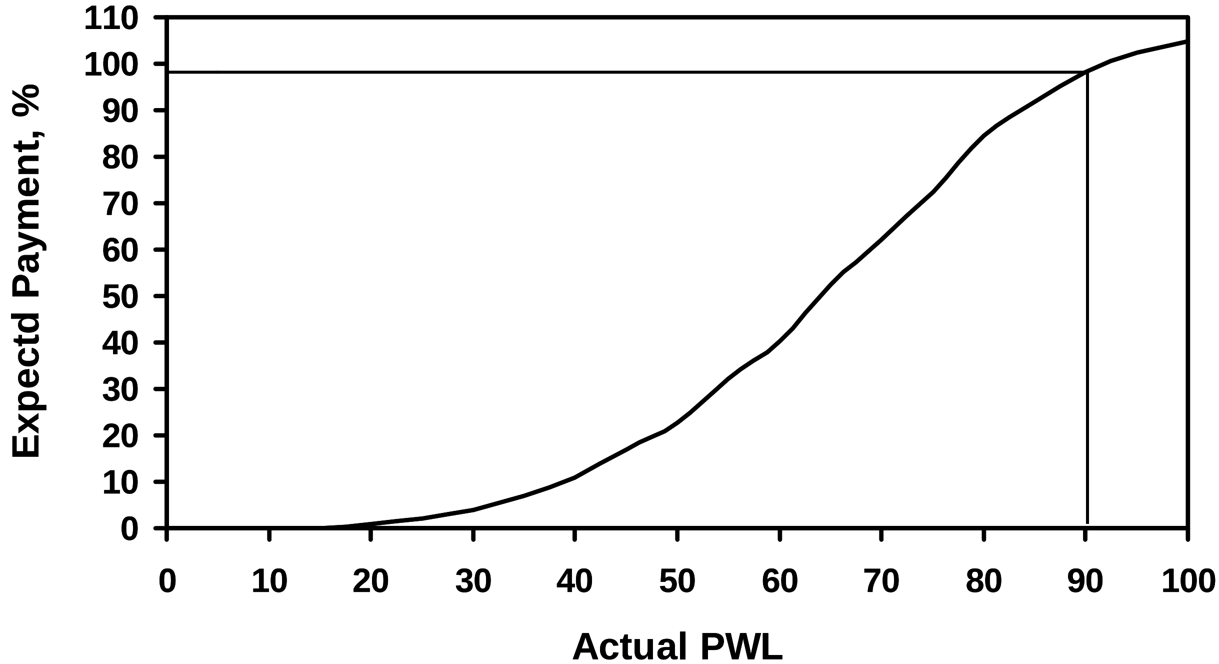


Figure 77. EP curve for the payment relationship Pay = 55 + 0.5*PWL*, with an RQL provision, sample size = 4.

For the plot in figure 77, the reason that the EP did not equal 100 at the AQL can be attributed to the provision requiring zero payment for estimated PWL values of less than 60. If the payment equation were allowed to apply throughout its total range, then the EP would have been able to reach 100 at the AQL. The reason that this did not happen can be shown by looking at the distribution of the estimated PWL values and their corresponding payment factors (shown in the histograms in figure 78). The histograms show that because of the variability of the small sample size, there were a number of times when the estimated PWL value for a lot was less than 60 even though the population PWL was 90. Similarly, there were many times when the estimated PWL was as high as 100.

When the estimated PWL was less than 60, a payment factor of zero was used to represent the remove-and-replace provision. This resulted in a number of zero payments for lots that are shown at the far left side of the payment histogram. Had these PWL values not been assigned as zero, the EP value could have been 100 at the AQL. To show this, the EP curve was calculated again without the remove-and-replace provision. This EP curve is shown in figure 79. The EP at the AQL is 100 percent and the EP curve exactly follows the payment equation.

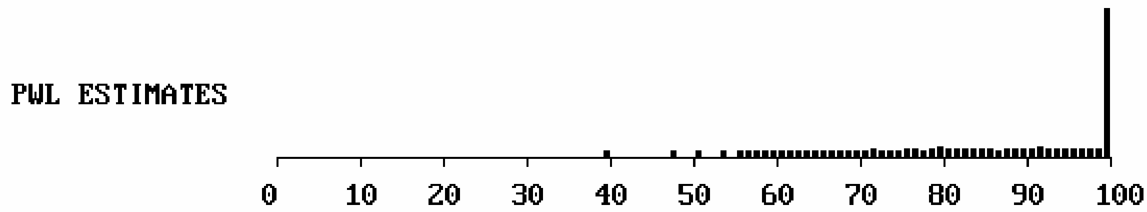


Figure 78a. Distribution of estimated PWL values for an AQL population.

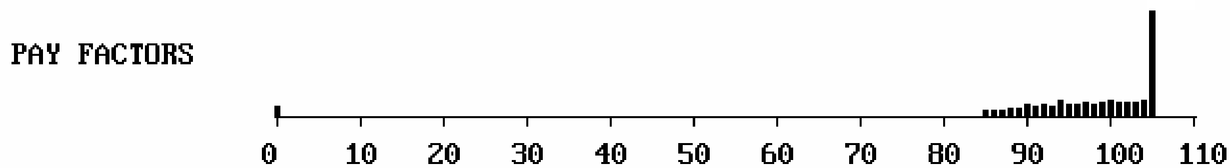


Figure 78b. Distribution of payment factors for an AQL population.

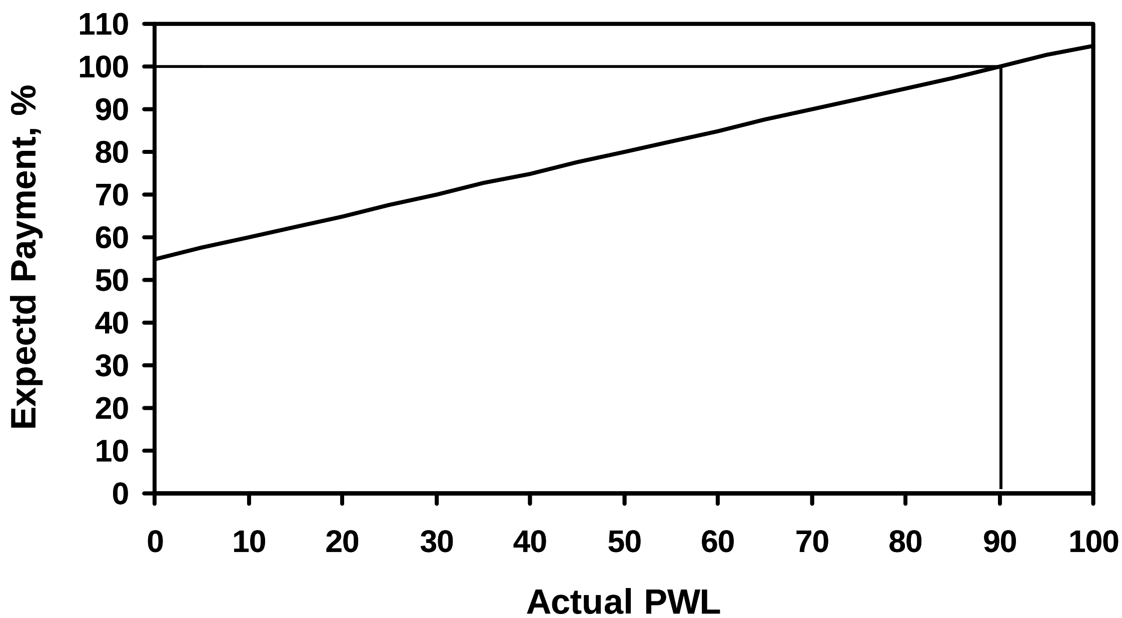


Figure 79. EP curve for the payment relationship Pay = 55 + 0.5*PWL*, sample size = 4.

The most common reason that the EP curve deviates from the payment equation is when there is no provision to allow for incentive, or bonus, payments. This, once again, creates a boundary at 100-percent payment. With no incentive payments, with a population at an AQL of 90 PWL, there will be times when the estimated PWL will be less than 90 and times when it will be greater than 90 (see the histogram in figure 78). When the estimated PWL is less than 90, a price reduction will be applied. When the estimated PWL is greater than 90, the payment will still be 100 percent if there is no incentive payment. There will be no payments greater than 100 percent to balance the payments less than 100 percent to allow the average payment (i.e., the EP) to be 100 percent.

Effect of Sample Size

As is shown in chapter 5, one way to reduce the variability of the individual estimated PWL values is to increase the size of the sample that is taken. This reduces the variability about the EP value and, thus, reduces the risks by increasing the likelihood that the estimated PWL values will be close to the true value.

The effect of sample size is not evaluated again in this chapter. However, the reduction in variability is shown very clearly in figures 9 through 12.

Distribution of Estimated PWL Values

Figure 79 clearly shows that even the maximum 5-percent bonus provided by equation 29 was sufficient to offset the times when the PWL for an AQL population is underestimated. This may not seem reasonable in light of the fact that equation 29 allows for only a 5-percent bonus, while allowing penalties that can theoretically reach as much as 45 percent. The reason that this can happen is related to the distribution of the estimated PWL values, particularly for the small sample sizes that are typically used in highway materials and construction.

The reason for this discrepancy is shown in figure 75 in the OC curve for the probability of receiving at least 100-percent payment. The OC curve shows that there is about a 40-percent chance of receiving less than 100-percent payment, while there is about a 60-percent chance of receiving 100-percent or greater payment. The fact that the sample PWL is an unbiased estimator of the population PWL does not mean that there is an equal chance of a high or low estimate. The fact that the average of the sampling distribution for individual PWL estimates has a mean equal to the true population PWL is what makes the estimator unbiased. In other words, the distribution of the estimates does not need to be symmetrical for the estimator to be unbiased.

It is the skewed distribution of sample PWL estimates that allows for a smaller bonus to offset larger penalties. The skewness of the distribution of the estimated PWL values and, hence, the estimated payment values can be shown with histograms.

Figure 80 shows histograms of the distribution of sample PWL estimates for a population with actual PWL = 90 and sample size = 4 for one-sided and two-sided specification limits. The histograms show that the distribution is skewed to the right for both one- and two-sided specifications, and that about 60 percent of the values lie above the actual PWL of 90.

Figure 81 shows similar histograms for a population with an actual PWL of 50. The two-sided specification exhibits the same skewness to the right, but with 60 percent below the actual PWL of 50. However, for the one-sided specification limit, the distribution appears to be symmetrical, with about 50 percent on either side of the actual PWL of 50.

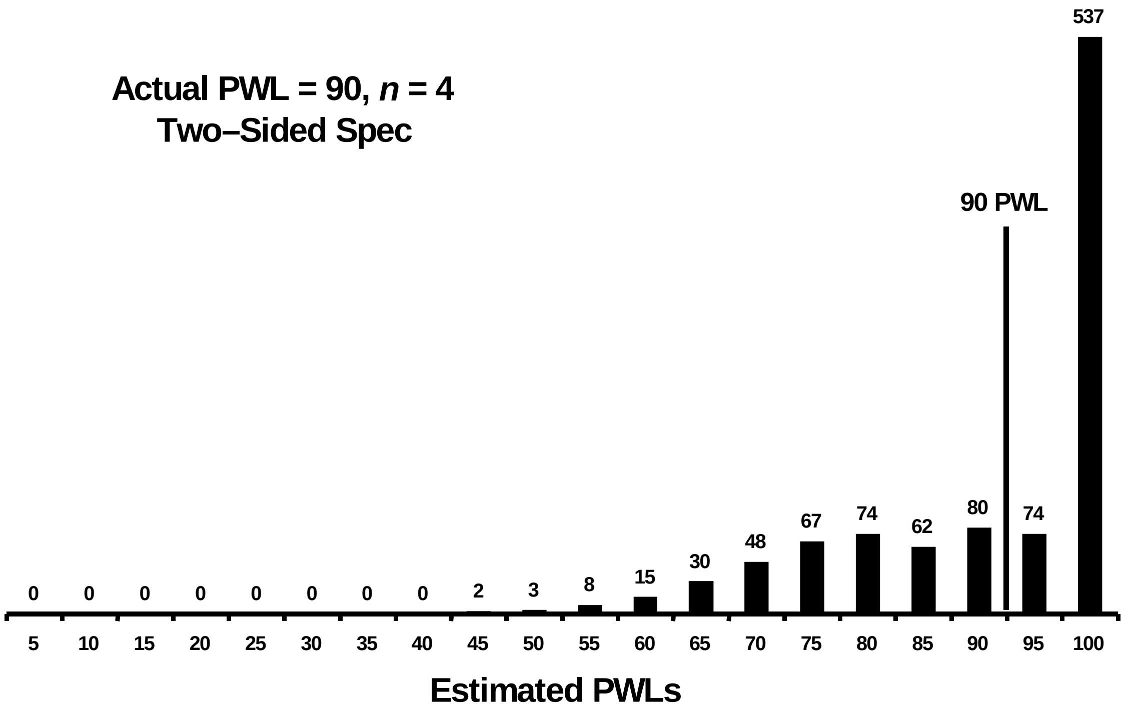
The reason that the two distributions are different lies in the fact that it is the standard deviation that causes the estimated PWL distribution to be skewed. The only way for a one-sided specification to have 50 PWL is if the population mean is centered on the specification limit. Any population that is centered on the one specification limit will have 50 PWL, regardless of the value of its standard deviation.

To illustrate why it is the standard deviation that causes the skewness in PWL estimates, it is necessary to look at the sampling distributions for sample means and sample standard deviations. It is well known that for any sample size, the distribution of the sample means from a normal distribution will be normally distributed with the mean equal to the population mean. This indicates that the sample mean should lead to a symmetrical distribution of sample PWL values.

However, it is also known that the sample variance follows a chi-square distribution. The chi-square distribution is skewed to the right, but the amount of skewness decreases as the sample size increases. Since the sample standard deviation is the square root of the sample variance, the distribution of the sample standard deviations will also be skewed. This is shown in figure 82, which shows the distribution of 1000 sample standard deviation values for sample sizes = 3, 5, and 10. Note that the spread of the values becomes less and the shape of the distribution approaches symmetry as the sample size increases.

Appendix H presents a more detailed discussion regarding the distributions of sample PWL estimates and how these distributions vary depending on the population PWL value.

Figure 80. Distributions of sample PWL estimates for a population with 90 PWL.



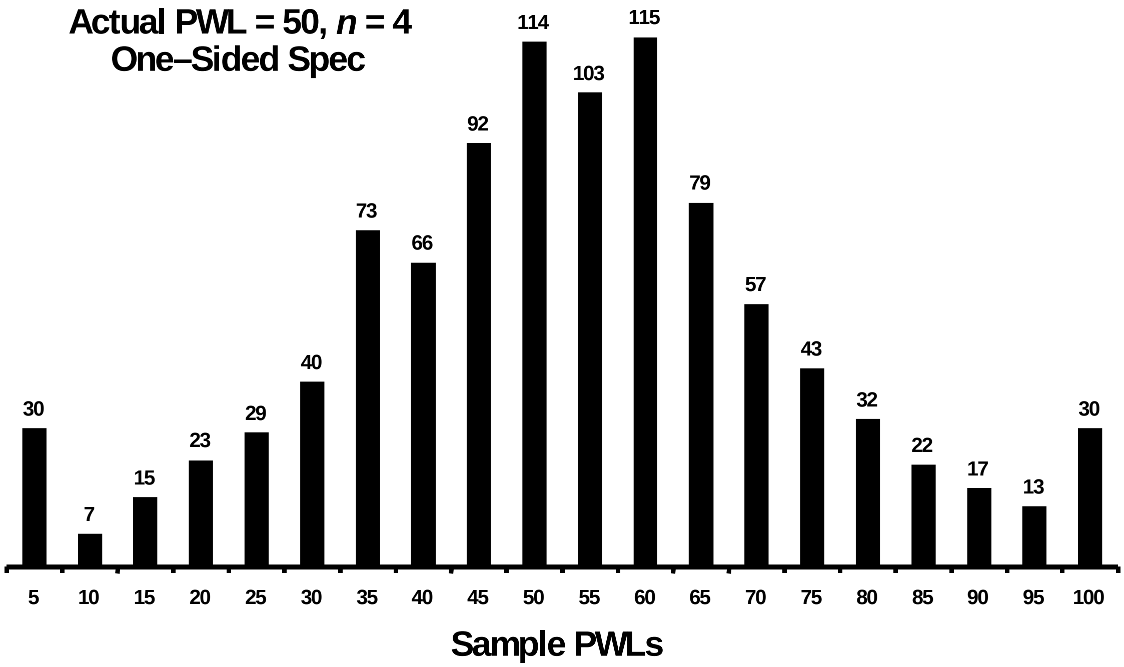


Figure 81a. Distributions of sample PWL estimates for a population with 50 PWL and one-sided speculations.

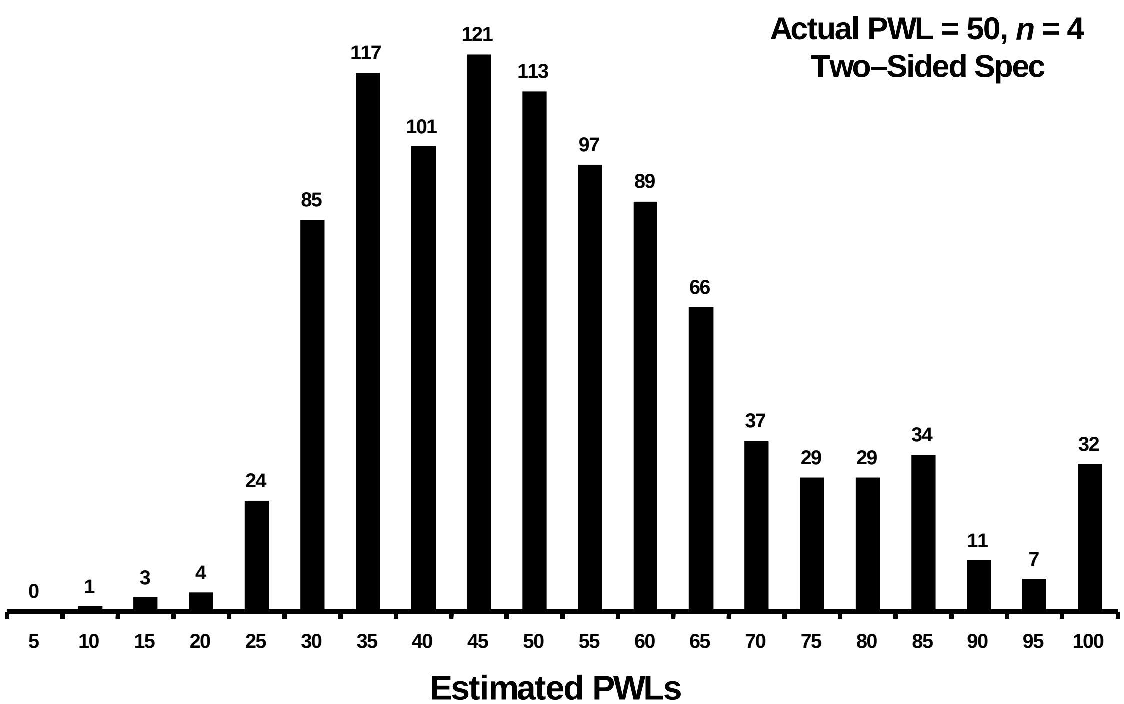


Figure 81b. Distributions of sample PWL estimates for a population with 50 PWL and two-sided speculations.

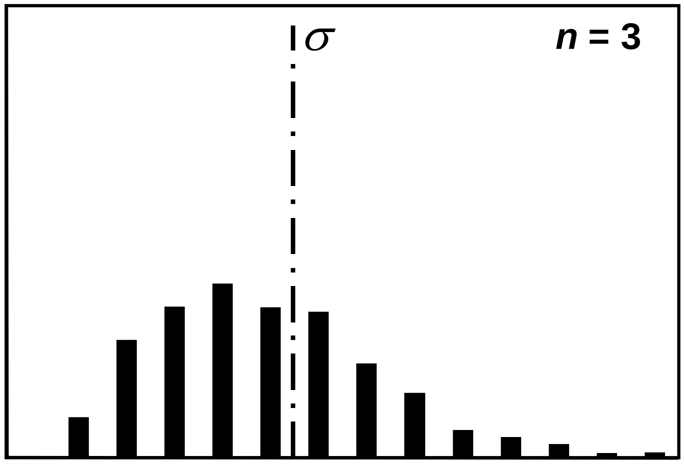


Figure 82a. Distribution of sample standard deviations for a sample size, n = 3, based on 1000 simulated samples.

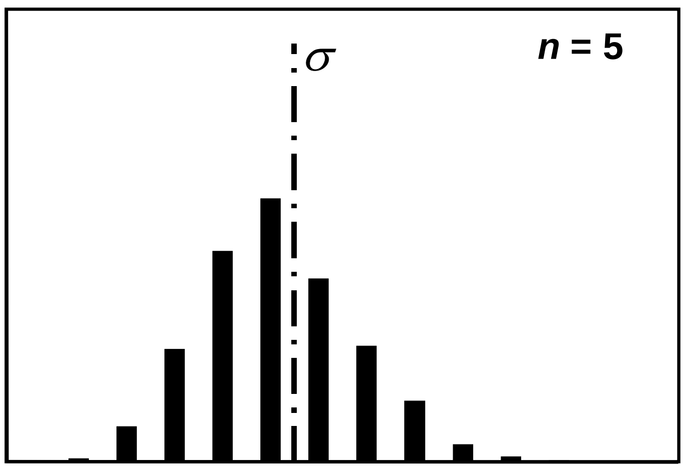


Figure 82b. Distribution of sample standard deviations for a sample size, n = 5, based on 1000 simulated samples.

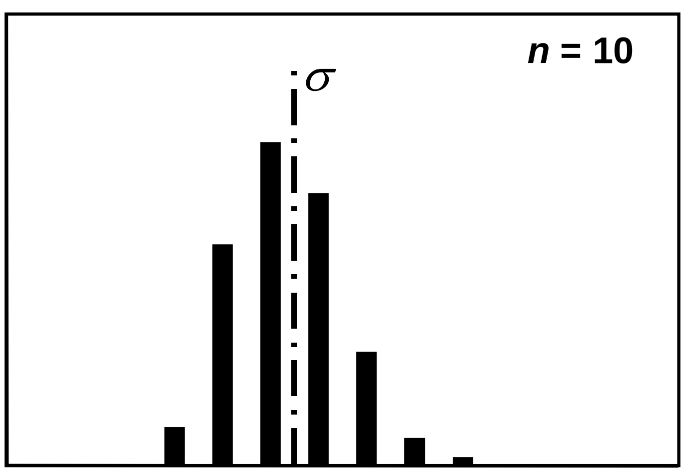


Figure 82c. Distribution of sample standard deviations for a sample size, n = 10, based on 1000 simulated samples.

CALCULATING RISKS FOR MULTIPLE QUALITY CHARACTERISTICS

The procedures presented above for OC and EP curves are primarily for the case of acceptance based on a single property. When, as will often be the case, there are multiple acceptance properties, it will be necessary for the agency to develop sophisticated computer simulation methods to complete a full analysis of the risks. These analyses will be quite involved and will be dependent on the quality characteristics chosen for acceptance and whether or not a performance model for predicting service life has been adopted by the agency. Another factor that will impact the analysis is whether a composite quality measure has been developed, or whether the individual quality measures will in some way, perhaps by adding, multiplying, or averaging, be combined into a composite payment factor. The topics of performance models and composite quality measures are addressed in chapter 11.

To illustrate the difficulty of trying to evaluate risks when there is more than one acceptance property, an example will be shown based on the use of two acceptance characteristics. The problem is simpler if it can be assumed that the characteristics are independent and, therefore, have a correlation coefficient of 0. If this might not be the case, then it must be determined whether or not it is necessary to develop separate EP contours or surfaces for each different possible set of correlation coefficients.

Calculating EP Contours

The same simulation routines that were presented in chapter 8 for evaluating two correlated quality characteristics can be used to develop a form of a two-variable EP curve. In chapter 8, it was shown that the EPs for two correlated variables did not change as the weighting factors or the value of their correlation coefficients varied. However, although the EPs did not vary, the standard deviations of the individual combined payment factors did vary with both the correlation coefficient and the weightings used when combining the individual payment factors.

Changes in the weighting factors are not likely to be a problem since they are usually part of the acceptance plan and, therefore, remain constant. The agency would have to decide whether or not it wished to evaluate the variability of the combined payment factors, or whether it was sufficient that the EP values matched the payment equations (i.e., that they were unbiased). If it is decided to only consider the EP, then a method needs to be developed to present the EP values in a way that can be used to assess the payment risks involved. One way to do this is by using contours or surfaces to represent the EP values as the actual PWL values for both acceptance characteristics vary. If only EP is considered, then the two variable EP contours or surfaces should not change with changes in the correlation coefficients.

To illustrate how this approach might be used to develop EP curves or surfaces, 1000 samples of size = 5 were simulated for correlation coefficients of +0.5 and –0.5, using several different methods for combining the individual payment factors into a combined payment factor. The simulations were conducted with actual PWL values ranging from 95 to 10 for both populations. The results of some of these simulation analyses are presented here for illustration.

Table 52 shows the results of simulations where the combined payments were determined as the average of the two individual payments. The individual payment factors were calculated using equation 29. The results show that there is no difference in the EP values when the correlation coefficient is +0.75, 0.00, or –0.75. The *correct* values for any cell in the table are determined by inserting the actual PWL values for variables 1 and 2 into equation 29. The EP values in the table are very close to what they should be to match the payment equation.

To illustrate how the EP values for two variables could be shown graphically, figure 83 shows EP contours for the EP values shown in table 53. These EP values were determined by multiplying the two individual payments, which were each obtained from equation 29. The values were based on a correlation coefficient of +0.50. Figure 84 shows another way in which the EP values could be represented as a surface in a three-dimensional plot. Still another two-dimensional approach for presenting the EP values for two variables is shown in figure 85. In this figure, the values in table 53 are plotted with one variable on the horizontal axis and the EP values on the vertical axis. The second variable is then plotted with a separate curve for each of a number of different PWL values.

While the three-dimensional approach is theoretically a good way to visualize how the EP values vary with the population PWL values, the two-dimensional methods are probably a more practical way to present the information. None of these visualization approaches would work if the number of acceptance variables were three or greater.

EFFECT OF ADDED PROVISIONS

From the example above, it may appear that the evaluation of the risks can be complicated. The complexity becomes much greater when provisions in addition to a simple payment equation are added to the acceptance plan. In the past, a popular provision has been to state that even if the estimated PWL indicates a payment reduction, no price reduction will be applied if all of the individual test results are within the specification limits. This would not only have the effect of raising the EP values, it would also make it much more complicated to simulate the specification to establish the EP values.

Another provision that might require analyses to fully understand the potential risks is a remove-and-replace, or leave in at zero payment, provision. These are the most extreme penalties that can be imposed; thus, it is incumbent upon the specifying agency to fully evaluate the risks for the contractor before implementing such a provision. It is relatively easy to evaluate the risk that removal and replacement will be required when there is a single acceptance quality characteristic. It merely requires the same approach that was used to determine the OC curves in figures 74 through 76. This is illustrated with a simple example.

Table 52. EP values using Pay = 55 + 0.5*PWL* for two individual payment factors and then averaging them, sample size = 5.

PWL Variable 1	PWL Variable 2							
	10	30	50	60	70	80	90	95
Correlation Coefficient = +0.75								
10	60.0	64.8	69.9	72.5	75.1	77.8	80.0	81.2
30	64.8	70.0	75.0	77.3	80.0	82.6	85.1	86.3
50	70.0	74.9	80.0	82.5	84.7	87.5	90.2	91.3
60	72.5	77.7	82.2	85.0	87.7	89.9	92.5	93.5
70	75.3	80.0	84.8	87.4	90.5	92.8	95.2	96.2
80	77.4	82.6	87.4	90.5	92.3	95.0	97.4	98.7
90	80.1	85.1	90.0	92.6	95.2	97.5	100.1	101.2
95	81.2	86.3	91.5	94.0	96.2	98.9	101.2	102.5
Correlation Coefficient = 0.00								
10	59.9	64.8	70.2	72.5	74.9	77.4	79.7	81.2
30	65.0	70.0	75.3	77.3	80.1	82.3	85.2	85.9
50	69.9	75.0	79.8	82.6	85.1	87.3	89.9	91.4
60	72.8	77.3	82.6	85.2	87.8	89.9	92.6	93.7
70	74.9	80.2	85.5	87.4	90.1	92.3	95.1	96.4
80	77.6	82.7	87.4	89.8	92.6	95.1	97.4	98.8
90	80.0	84.8	89.6	92.6	94.9	97.7	99.7	101.2
95	81.3	86.0	91.3	93.6	96.4	98.9	101.1	102.5
Correlation Coefficient = –0.75								
10	59.9	65.0	70.2	72.6	74.9	77.5	80.0	81.1
30	64.9	69.7	75.0	77.6	79.9	82.4	85.0	86.0
50	69.8	75.0	79.9	82.6	85.1	87.6	90.3	91.1
60	72.5	77.4	82.5	84.9	87.6	89.8	92.5	94.0
70	74.9	80.2	84.9	87.6	90.0	92.4	94.7	96.2
80	77.5	82.6	87.5	90.2	92.8	95.0	97.6	98.8
90	79.9	84.9	90.2	92.5	94.8	97.4	100.1	101.2
95	81.2	86.2	91.3	94.0	96.4	98.9	101.0	102.6

Table 53. EP values using Pay = 55 + 0.5*PWL* for two individual payment factors and then multiplying them, sample size = 5, correlation coefficient = +0.5.

PWL Variable 1	PWL Variable 2							
	10	**30**	**50**	**60**	**70**	**80**	**90**	**95**
10	35.9	42.1	48.2	50.6	54.2	56.8	60.2	61.3
30	42.1	49.1	56.5	59.8	62.8	66.4	69.8	71.4
50	48.9	56.7	64.7	68.6	72.7	76.6	80.6	81.6
60	51.2	59.5	68.4	72.9	77.3	80.7	85.2	87.0
70	54.4	63.2	72.4	77.3	81.5	85.8	90.0	92.3
80	57.5	66.8	75.7	80.9	86.0	90.9	95.2	97.5
90	59.9	70.1	80.5	84.9	90.6	95.1	99.7	102.0
95	61.4	72.2	82.2	86.7	92.7	97.6	102.9	105.0

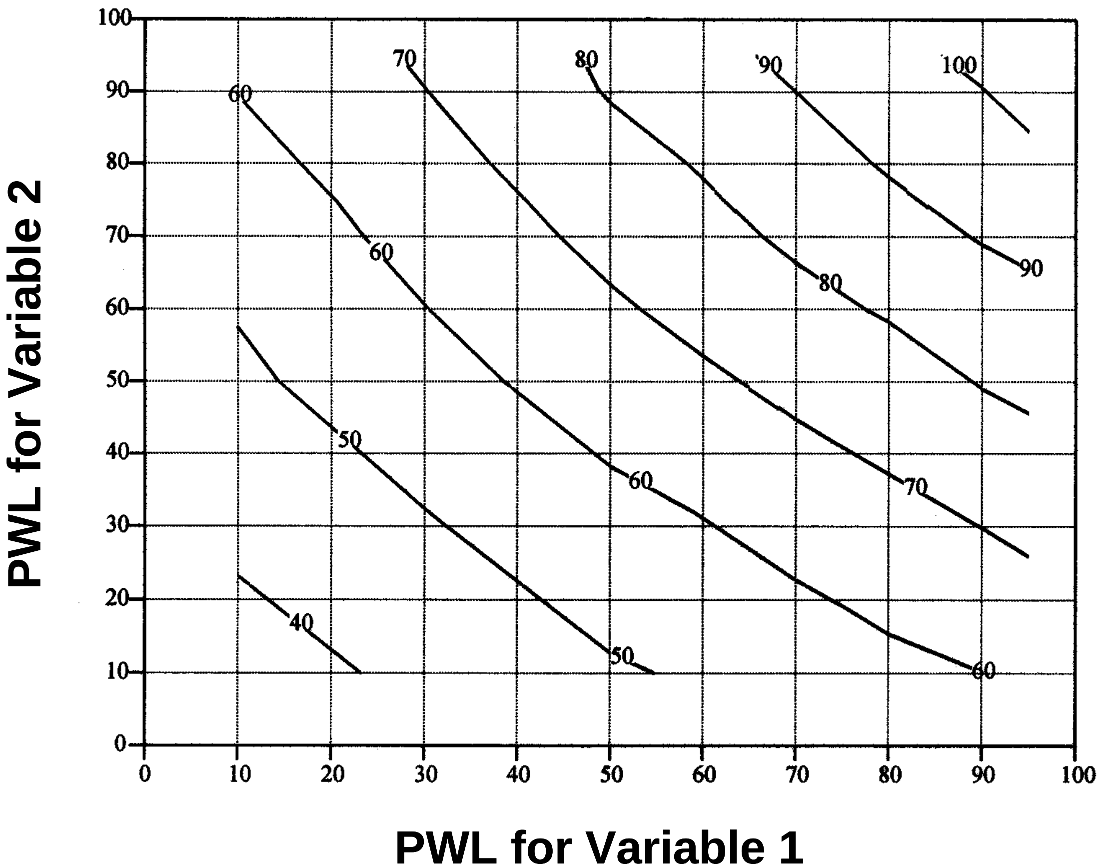


Figure 83. EP contours for the values in table 53.

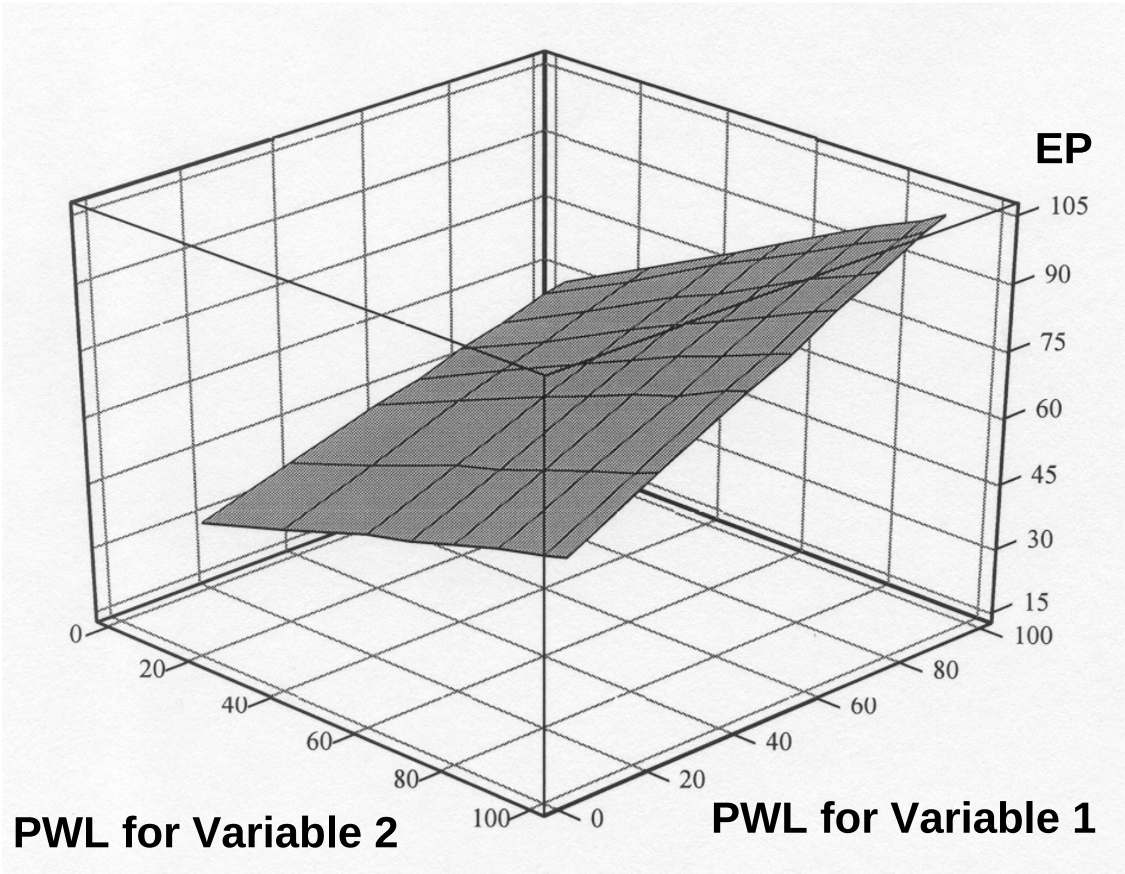


Figure 84. EP surface for the values in table 53.

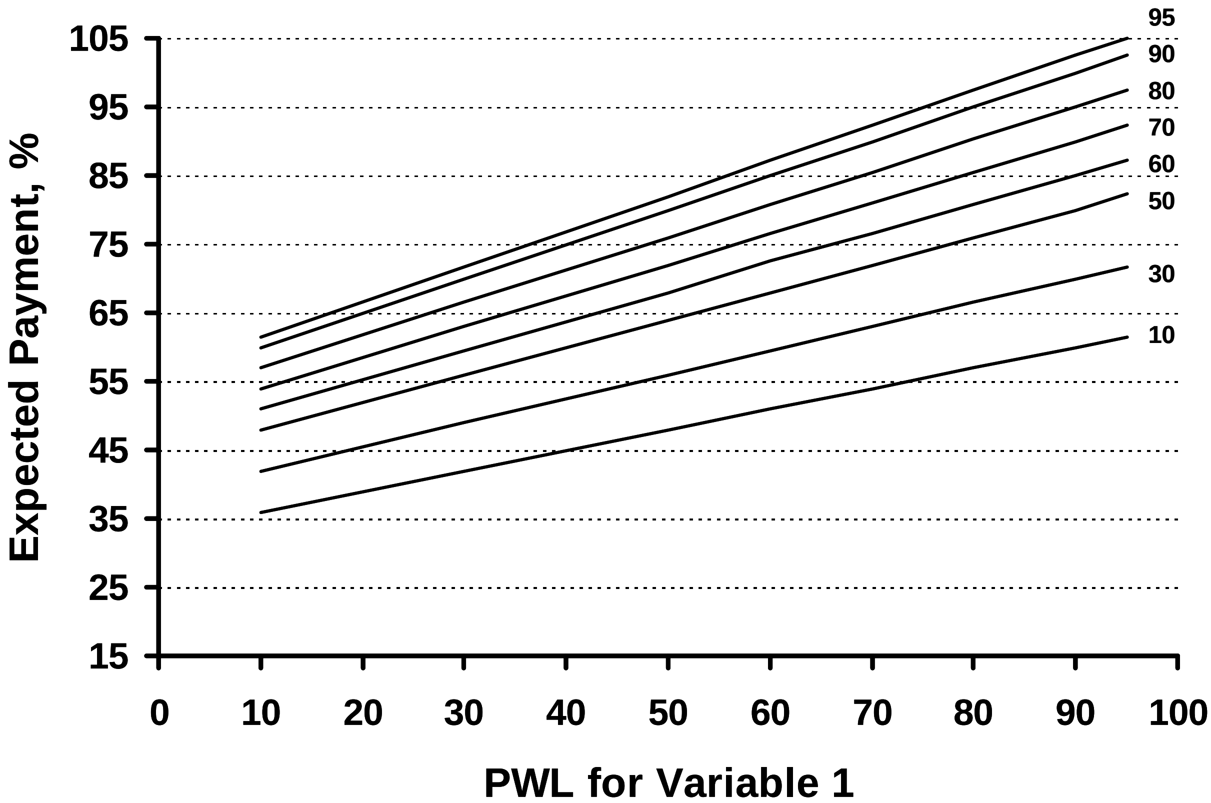


Figure 85. EP curves for the values in table 53.

Suppose that an acceptance plan calls for a sample size = 5 and requires that the material be removed and replaced if the estimated PWL value is less than 60. The OC curve for the probability that removal and replacement will be required is determined by the probability that a population with any given PWL value will yield a sample PWL of less than 60. The OC curve (the probability of acceptance) associated with the remove-and-replace decision is shown in figure 86. This OC curve represents the probability that the material will NOT require removal and replacement. The probability that removal and replacement would be required would be 1.00 minus the value indicated on the vertical axis.

Since small sample sizes, usually n = 3-5, are often involved in the acceptance decision, there is a large amount of variability in the estimated PWL value for any given population. Table 54 shows the probability that populations of various quality levels would have an estimated PWL of less than 60 for a single quality characteristic and would thus require removal and replacement.

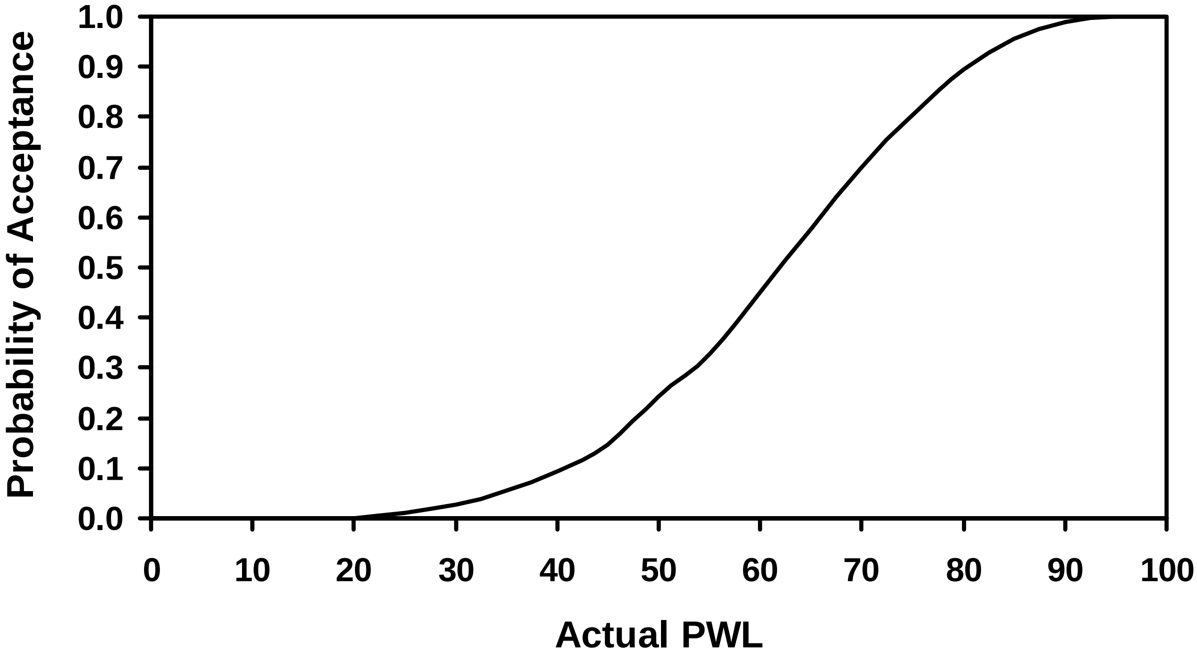


Figure 86. OC curve for an acceptance plan that calls for rejection if the estimated PWL is less than 60, sample size = 5.

Suppose that the agency had selected 90 PWL as the AQL for the acceptance plan. Table 54 shows that when the sample size = 3 for an AQL population (i.e., 90 PWL), each single quality characteristic has more than a 5-percent chance of yielding a sample result that indicates that removal and replacement are required. For sample size = 5, this probability drops to about 1 percent.

To further complicate the situation, many acceptance plans call for up to four or more acceptance characteristics. If the same remove-and-replace provision is applied if any of the acceptance characteristics has an estimated PWL of less than 60, then the risks are considerably greater. Assume that our example acceptance plan has four acceptance characteristics and requires removal and replacement if any of them have estimated PWL values of less than 60. If we assume that the four characteristics are independent, table 54 shows the probability that at least one of the four characteristics will trigger the remove-and-replace provision.

For a sample size = 3, there is about a 20-percent chance that an AQL population will require removal and replacement. Even for a sample size = 5, there is more than a 4-percent chance of the need for removal and replacement.

Table 54. Probability that populations with various quality levels would require removal and replacement for the example in figure 86.

Population PWL	With One Quality Characteristic			With Four *Independent* Quality Characteristics		
	***n* = 3**	***n* = 4**	***n* = 5**	***n* = 3**	***n* = 4**	***n* = 5**
100	0.000	0.000	0.000	0.000	0.000	0.000
95	0.014	0.003	0.001	0.055	0.012	0.004
90	0.054	0.027	0.011	0.199	0.104	0.043
85	0.121	0.068	0.044	0.403	0.245	0.165
80	0.199	0.146	0.106	0.588	0.468	0.361
75	0.285	0.240	0.197	0.739	0.666	0.584
70	0.366	0.339	0.302	0.838	0.809	0.763
65	0.471	0.449	0.423	0.922	0.908	0.889
60	0.563	0.555	0.549	0.964	0.961	0.959
55	0.631	0.651	0.670	0.981	0.985	0.988
50	0.710	0.735	0.758	0.993	0.995	0.997
45	0.790	0.822	0.854	0.998	0.999	1.000
40	0.841	0.887	0.906	0.999	1.000	1.000
35	0.894	0.927	0.944	1.000	1.000	1.000
30	0.938	0.961	0.971	1.000	1.000	1.000
25	0.958	0.979	0.988	1.000	1.000	1.000
20	0.981	0.993	1.000	1.000	1.000	1.000
15	1.000	1.000	1.000	1.000	1.000	1.000
10	1.000	1.000	1.000	1.000	1.000	1.000
5	1.000	1.000	1.000	0.000	1.000	1.000
0	1.000	1.000	1.000	0.055	1.000	1.000

CLOSING COMMENTS ON RISK

OC and EP curves describe the operation of an acceptance plan such that the risks can be evaluated throughout the entire quality regime. If the risks are considered to be acceptable, no modifications to the initial acceptance plan are necessary. However, if the risks are considered to be unacceptable in terms of being too high for either or both parties, a reassessment of the acceptance plan is necessary.

There is no easy answer to the question "Are the risks acceptable?" The decision regarding what does or does not constitute an acceptable level of risk will, to a great extent, be a subjective one. There is, however, one factor that is not subjective. There is generally universal agreement that the EP should be 100 percent for quality that is at exactly the AQL. Although it should not be confused with the statistical risk (α), the agency may wish to consider the *average payment* risk to the contractor if the EP is less than 100 percent at the AQL, or to the agency if the EP is greater than 100 percent at the AQL. The EP at the RQL is another point that is often specifically considered.

It must be remembered that the EP alone is not a complete measure, particularly the likelihood that any *individual* lot will receive a correct payment factor. The variability of the individual payment factors about the EP curve must also be considered. Ultimately, the decision regarding what constitutes acceptable or unacceptable risks rests with the individual agency.

10. THE COMPOSITE PAYMENT FACTOR

TRADITIONAL COMPOSITE PAYMENT FACTORS

Most specifications employ multiple quality characteristics for acceptance and payment. Therefore, one issue that must be addressed is how to combine these multiple characteristics to come up with a single payment factor for a lot. In chapter 8, a quality measure (PWL) was selected for use in an equation to determine payment for a single quality characteristic. A number of different methods for determining composite payment factors are currently in use. These can be referred to as *traditional* methods for determining a composite payment factor.

Through the years, various agencies have considered at least five different approaches for combining a number of payment factors for individual acceptance quality characteristics into a single composite payment factor. These approaches include:

- Averaging (possibly with weighting factors) the individual payment factors.
- Multiplying the individual payment factors.
- Summing the individual payment adjustments.
- Using the maximum individual payment factor.
- Using the minimum individual payment factor.

The approach of using the maximum individual payment factor could be thought of as allowing the contractor the benefit of the doubt. The approach using the minimum individual payment factor as the composite payment factor is based on the *weak link* theory (i.e., the lowest payment factor indicates the value for all of the quality characteristics). For the other three approaches, the concept is that all of the individual factors contribute to the total. However, the composite payment for the three approaches can be quite different depending on the value of the individual payment factors. All of these approaches for determining a composite payment factor were analyzed by computer simulation.

ANALYSES PROCEDURES

To simplify the simulation process and to allow for graphical presentation of the results, the analyses of the different methods for determining a composite payment factor were conducted only for the case where there are two individual, but possibly correlated, payment characteristics and where PWL was used for the quality measure. Only the case where the two individual quality characteristic populations have the same PWL value is presented here. All analyses were conducted using a sample size = 5, although similar trends would be expected for other sample sizes; however, the variability would be greater for smaller sample sizes and less for larger sample sizes.

The methods considered for combining the individual payment factors were:

- Averaging the individual payment factors $((0.5 \times P_1) + (0.5 \times P_2))$.
- Using a weighted average of the individual payment factors $((0.75 \times P_1) + (0.25 \times P_2))$.

- Multiplying the individual payment factors ($P_1 \times P_2$).
- Summing the individual payment adjustments ($P_1 + P_2 - 100$).
- Using the maximum individual payment factor (Max(P_1, P_2)).
- Using the minimum individual payment factor (Min(P_1, P_2)).

Since the two simulated populations had equal PWL values, it was anticipated that the average and weighted average methods would have similar results for the EP and would exhibit similar trends for standard deviation, although the magnitudes of the standard deviation values would be different because of the different weighting factor multipliers used.

Each of the above payment calculation methods was simulated for three different equal population PWL values (90, 70, and 50). Each of the PWL values was simulated with using correlation coefficients ranging from –1.0 to +1.0. The results of these analyses are presented below.

ANALYSES RESULTS

The results of the analyses are shown in figures 87 through 95. Figures 87 through 92 show the expected combined payment and standard deviation of the individual combined payments for correlation coefficients = –1.00, –0.90, –0.75, –0.50, –0.25, –0.10, 0.00, +0.10, +0.25, +0.50, +0.75, +0.90, and +1.00 for each of the six methods identified above. The simulation results for three different PWL values (90, 70, and 50) are included in each figure.

Figures 87 through 90 show similar trends regarding EP and standard deviation values for the simulated combined payments using the averaging, weighted average, multiplying, and summing methods. These figures show that the level of correlation between the two acceptance variables does not affect the EP values for these four methods. These figures also show that there is a similar effect on the standard deviation of the payment values for both different PWL values and for different levels of correlation. For the 50 and 70 PWL values, there is a consistently increasing trend in standard deviation as the correlation coefficient goes from –1.0 to +1.0. However, for the 90 PWL value, the standard deviations decrease slightly as the correlation coefficients approach 0.0 from both the positive and negative directions.

The reason for these different trends in the standard deviation results rests in the fact that the upper limit of 100 PWL (you cannot have more than 100 PWL) is reached much more often for a population of 90 PWL. This boundary has the effect of limiting the spread of the individual payment values, thereby reducing the standard deviation values. The discussions in chapter 8 that explained why the standard deviation was smaller for negative correlations and larger for positive correlations explain this similar trend in figures 87 through 90.

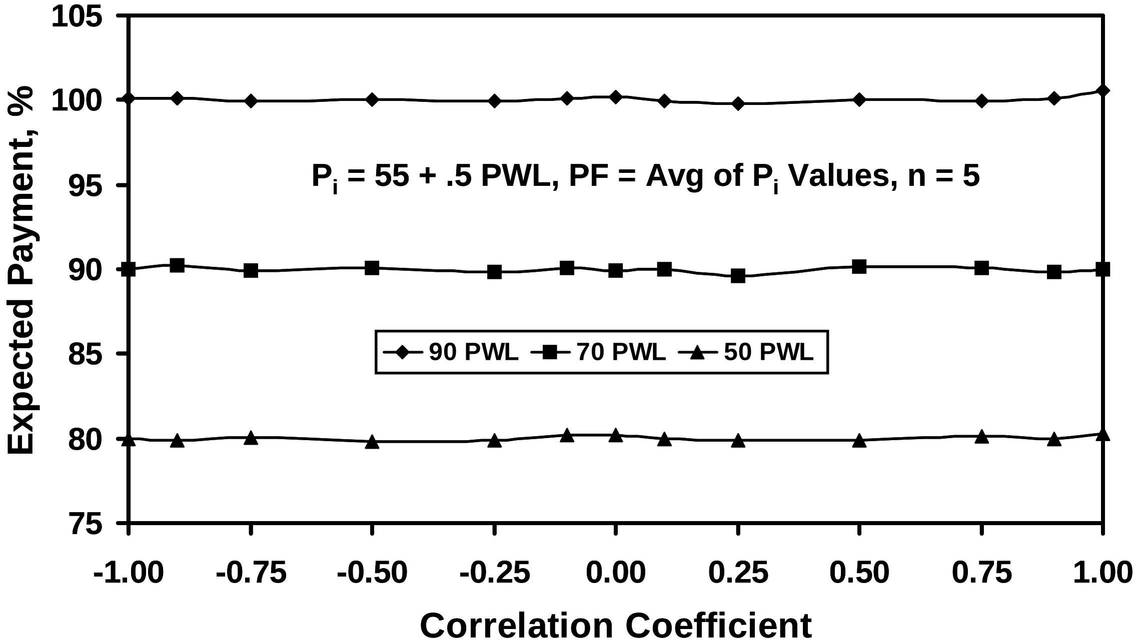


Figure 87a. Simulation results of expected payment for the averaging method, combining two populations with equal PWL values.

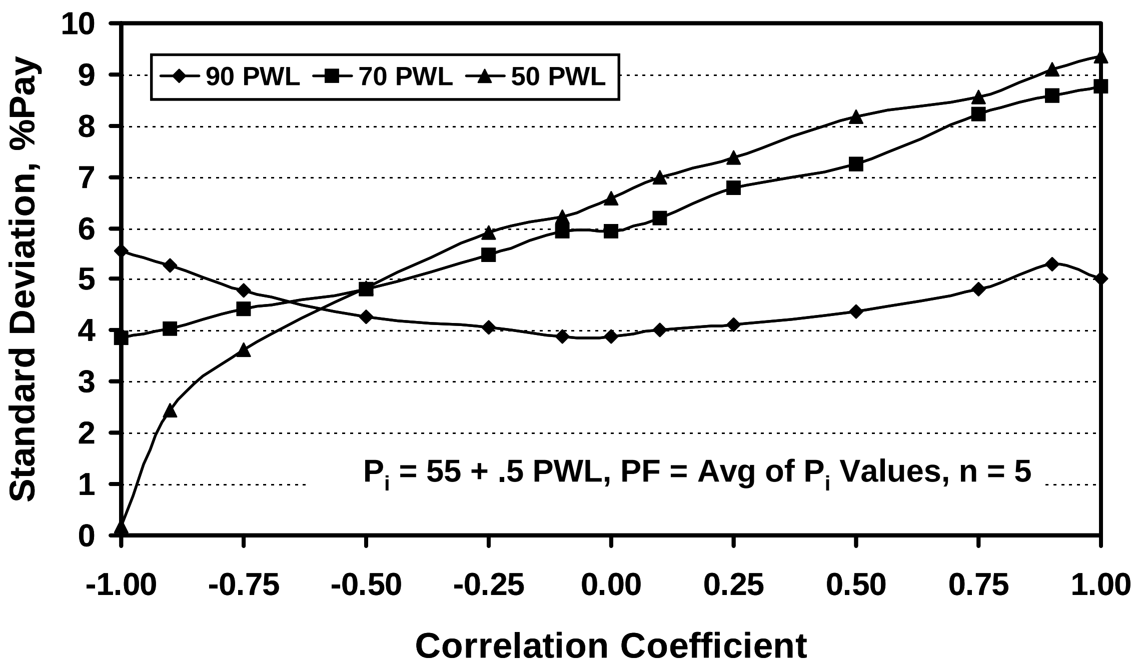


Figure 87b. Simulation results of standard deviation values for the averaging method, combining two populations with equal PWL values.

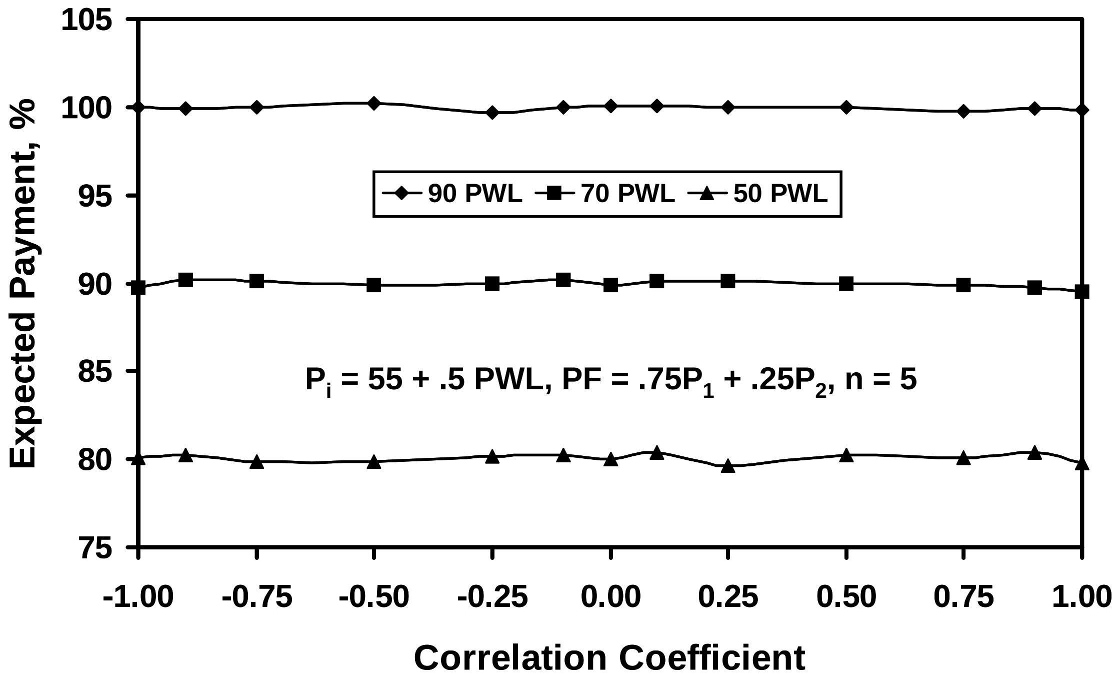


Figure 88a. Simulation results of expected payment for the weighted average method, combining two populations with equal PWL values.

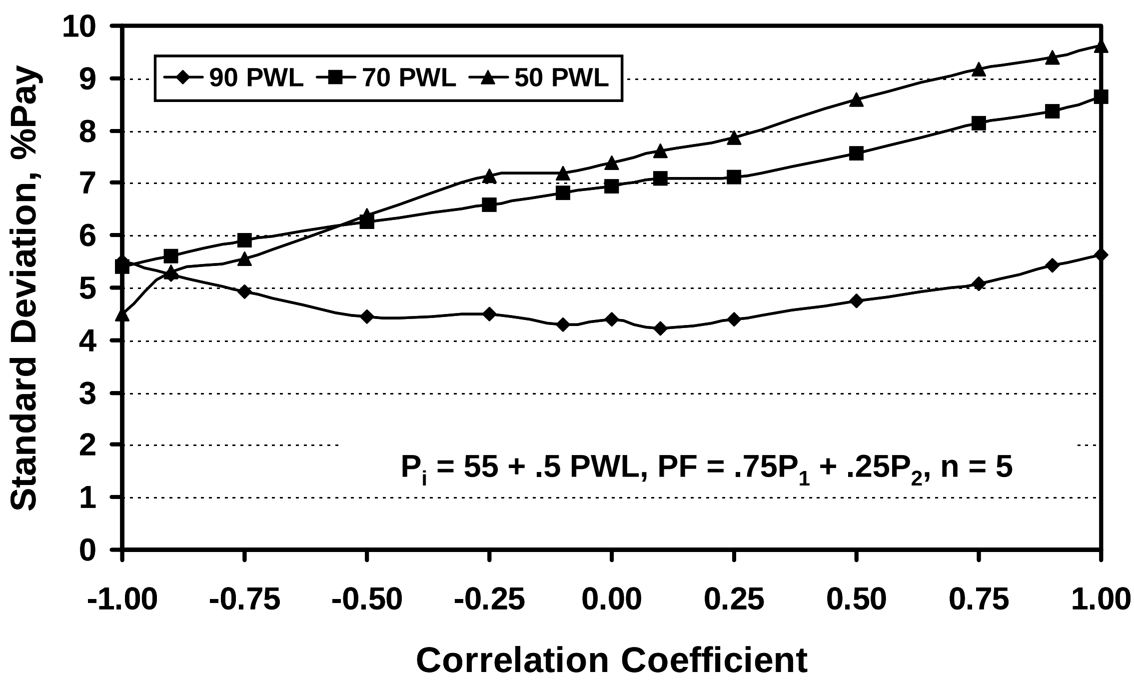


Figure 88b. Simulation results of standard deviation values for the weighted average method, combining two populations with equal PWL values.

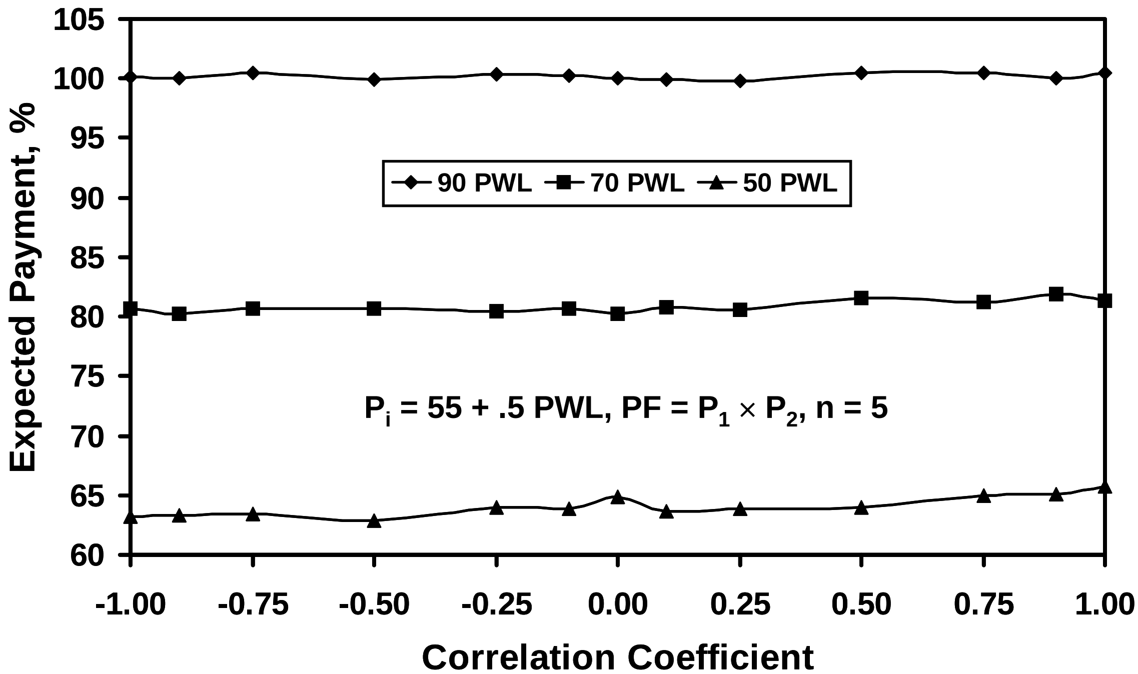


Figure 89a. Simulation results of expected payment for the multiplication method, combing two populations with equal PWL values.

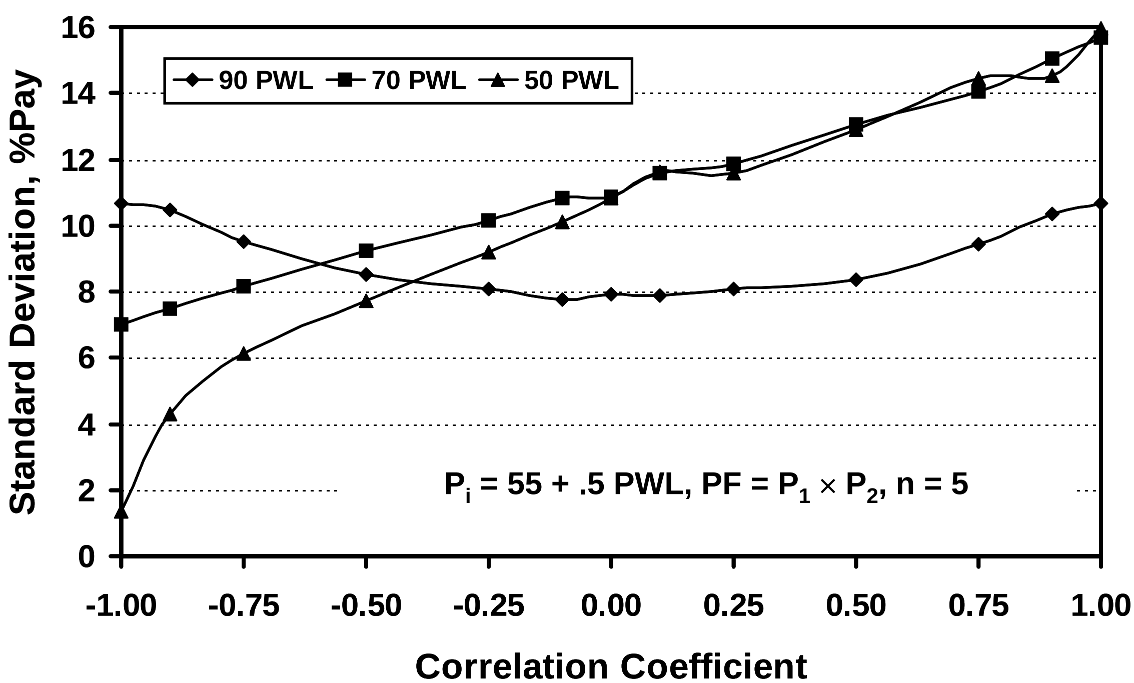


Figure 89b. Simulation results of standard deviation values for the multiplication method, combining two populations with equal PWL values.

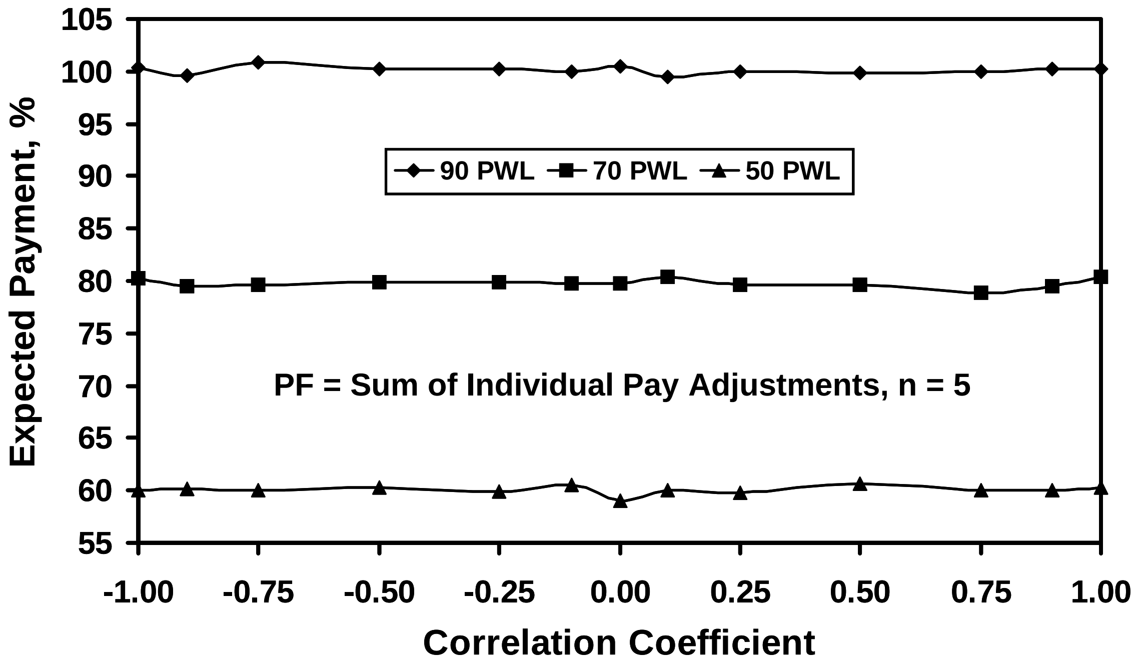


Figure 90a. Simulation results of expected payment for the summation method, combining two populations with equal PWL values.

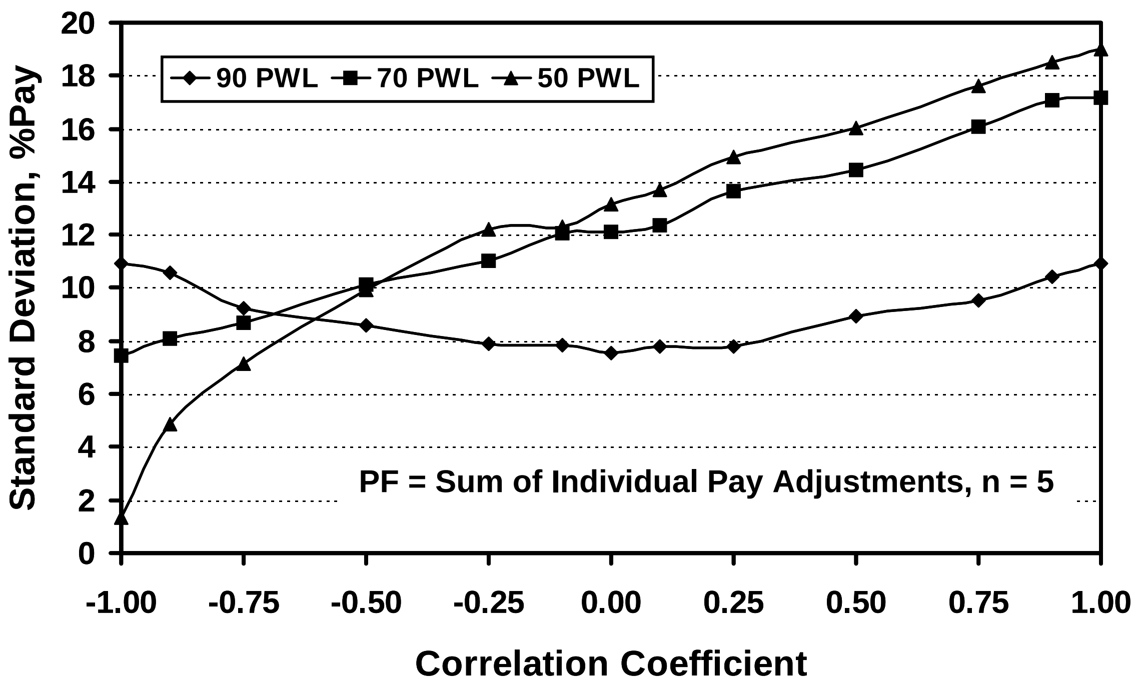
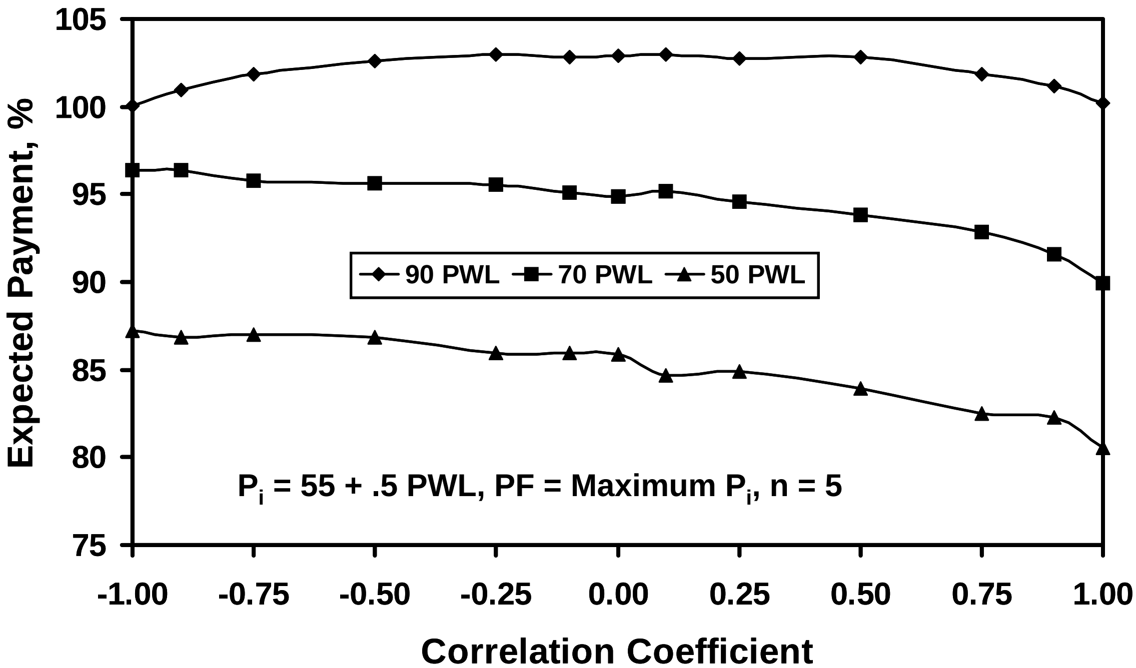


Figure 90b. Simulation results of standard deviation values for the summation method, combining two populations with equal PWL values.



91a. Simulation results of expected payment for the maximum method, combining two populations with equal PWL values.

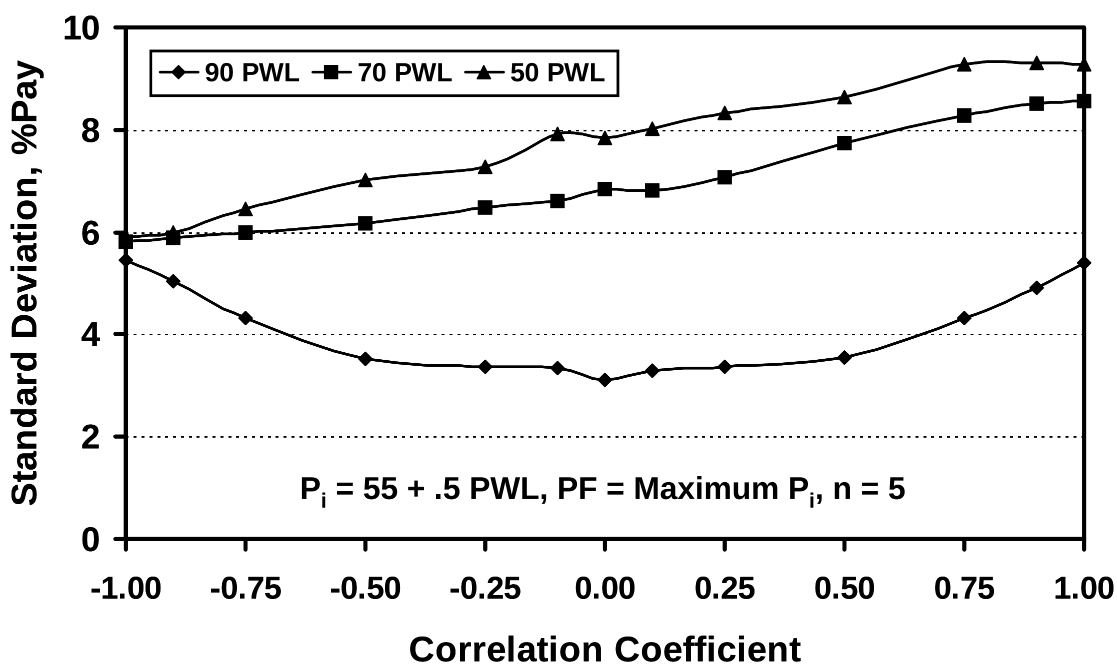


Figure 91b. Simulation results of standard deviation values for the maximum method, combining two populations with equal PWL values.

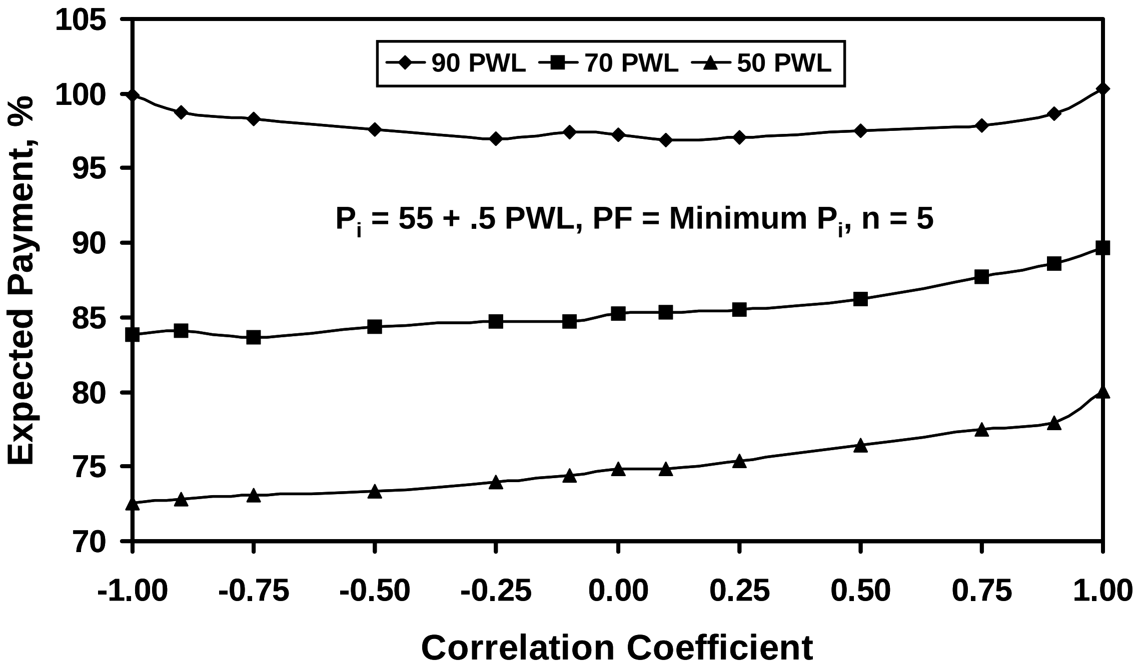


Figure 92a. Simulation results of expected payment for the minimum method, combining two populations with equal PWL values.

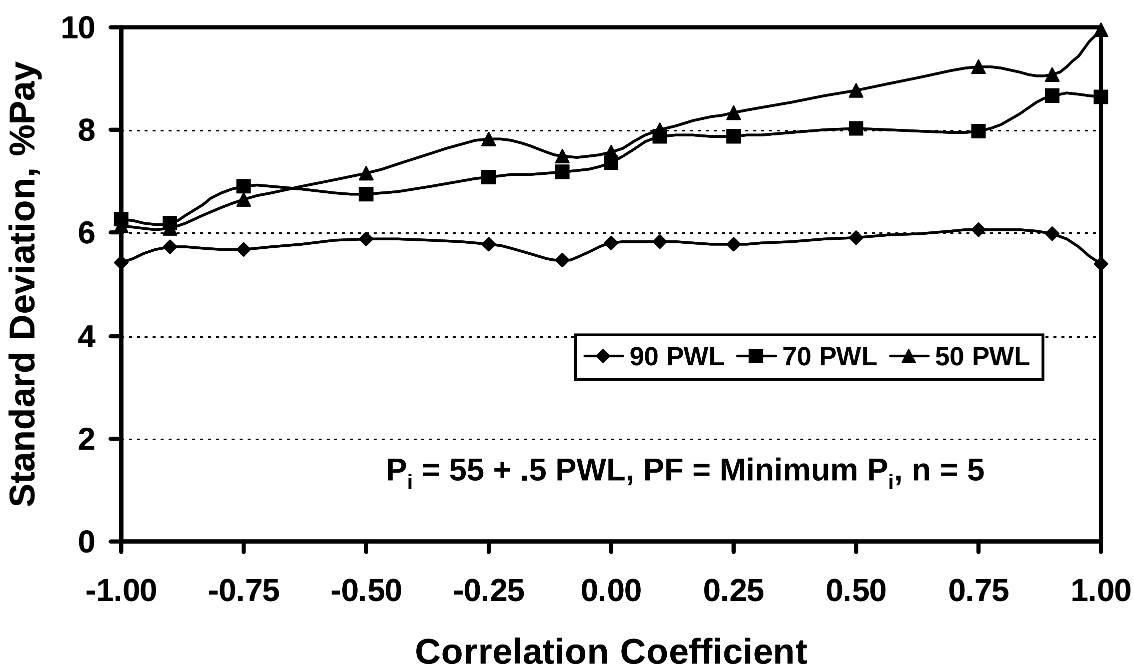


Figure 92b. Simulation results of standard deviation values for the minimum method, combining two populations with equal PWL values.

Somewhat different trends are shown in figures 91 and 92, which show the results for the maximum and minimum methods, respectively, for determining the combined payment factor. For these methods, for populations of 70 and 50 PWL, the EP values vary with the level of correlation between the two acceptance variables. This is not a desirable situation. In the first four methods considered, the EP for each level of quality, as indicated by the PWL, was independent of the correlation coefficient. For these methods, the variability was affected by the correlation value, but the EP was centered on the correct value.

In figure 91, not only does the variability increase as the correlation goes from –1.0 to +1.0, but the EP values for populations of 70 and 50 PWL decrease over this same range. In figure 92, the EP values for populations of 70 and 50 PWL increase as the correlation goes from –1.0 to +1.0. For the maximum and minimum methods, the variability increases and the increasing or decreasing EPs are related to one another. If the average value remains constant, but variability increases, then there is more spread about the average. Since it is the single largest value that establishes the payment for the maximum method, the increased variability means that the maximum values will, on average, be higher. Similarly, for the minimum method, the increased variability also means that the minimum values will be lower.

Comparing the Methods

Figures 93 through 95 show the same results as figures 87 through 92, but plotted by the actual PWL values rather than the payment combination method. In these figures, it is clear that the EP values for the averaging, weighted average, multiplying, and summing methods do not vary with the correlation between the acceptance variables, and that the maximum and minimum methods do vary.

When the EP values vary with the correlation coefficient, it means that the average payment values will be different in the long run for different correlation relationships between the two acceptance variables. When the EP values do not vary with the correlation coefficient, but the standard deviation does vary, this means that even though the average payment values in the long run will be the same, there will be a difference in the spread or range of payment values that may be estimated for individual lots.

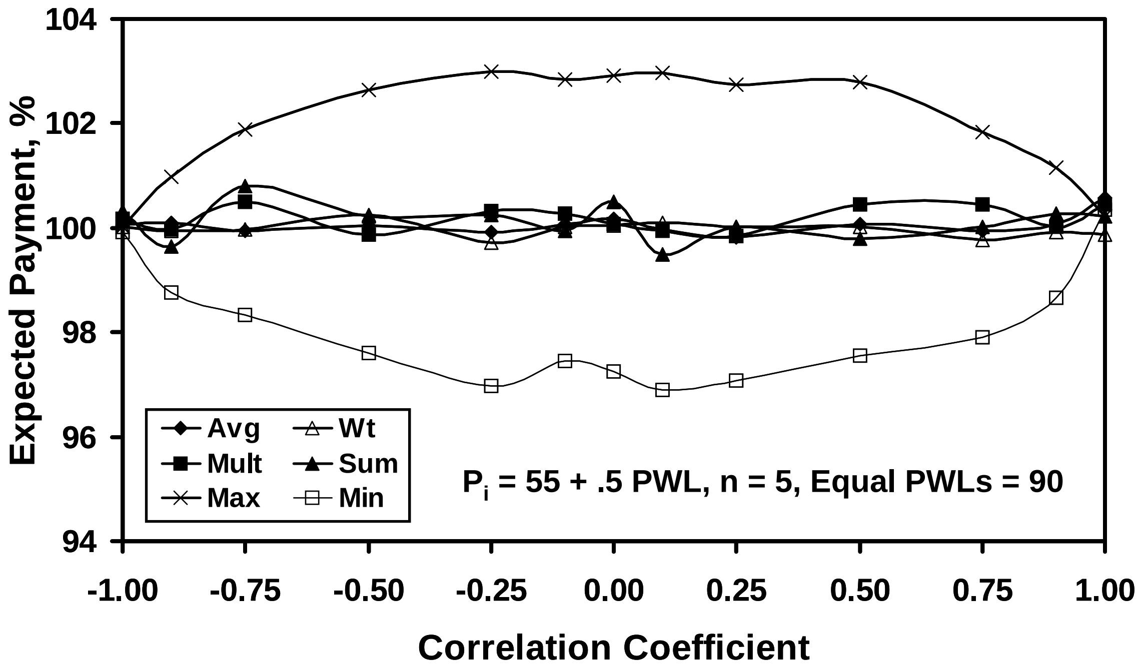


Figure 93a. Comparison of simulation results for various methods for combining individual expected payment factors for two populations with PWL = 90.

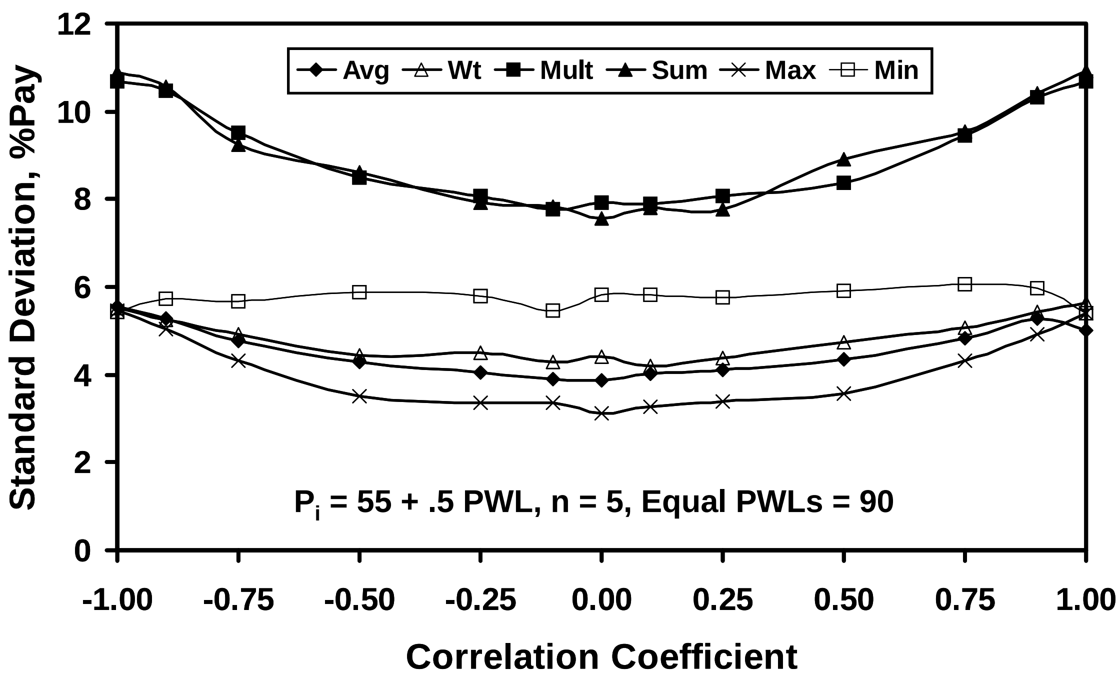


Figure 93b. Comparison of simulation results for various methods for combining individual standard deviation payment factors for two populations with PWL = 90.

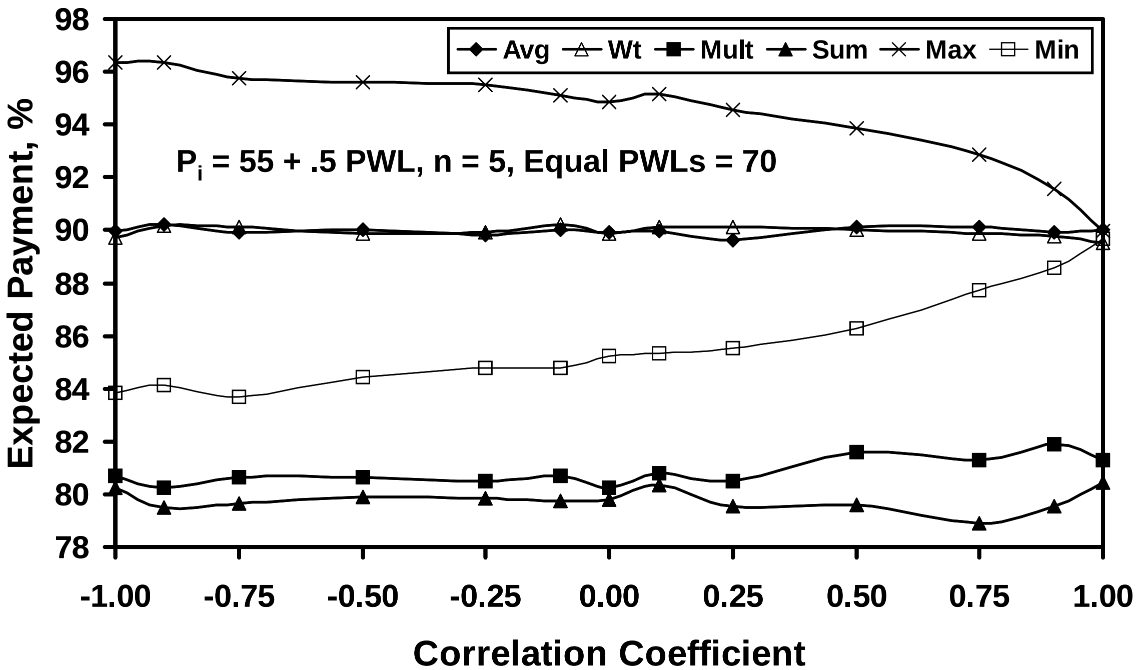


Figure 94a. Comparison of simulation results for various methods for combining individual expected payment factors for two populations with PWL = 70.

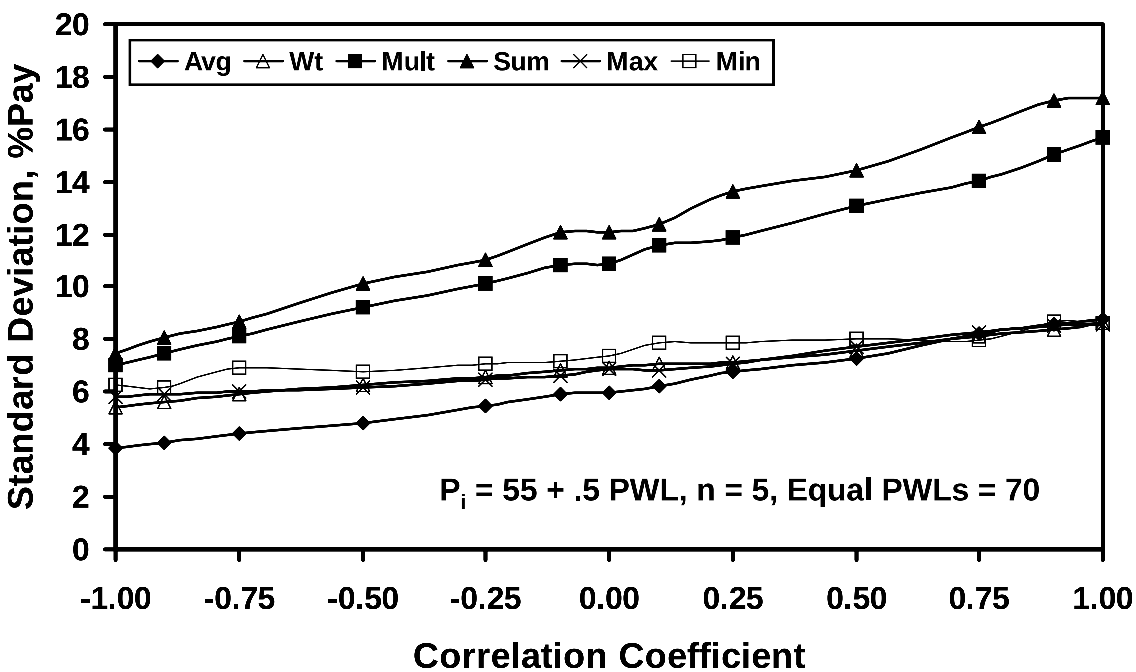


Figure 94b. Comparison of simulation results for various methods for combining individual standard deviation payment factors for two populations with PWL = 70.

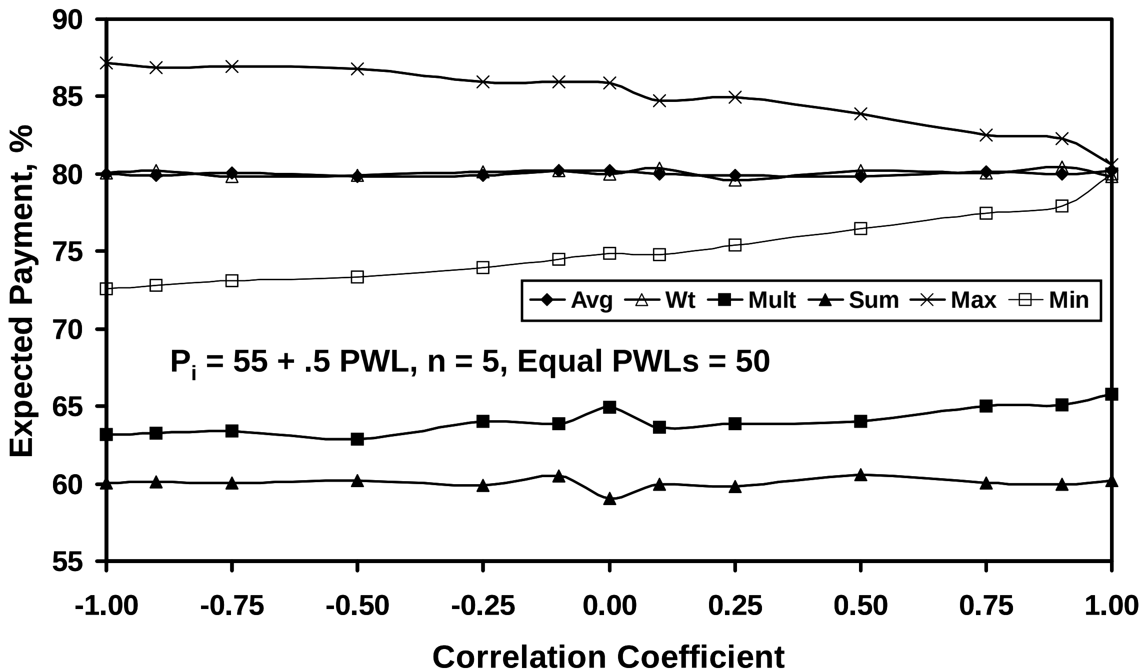


Figure 95a. Comparison of simulation results for various methods for combining individual expected payment factors for two populations with PWL = 50.

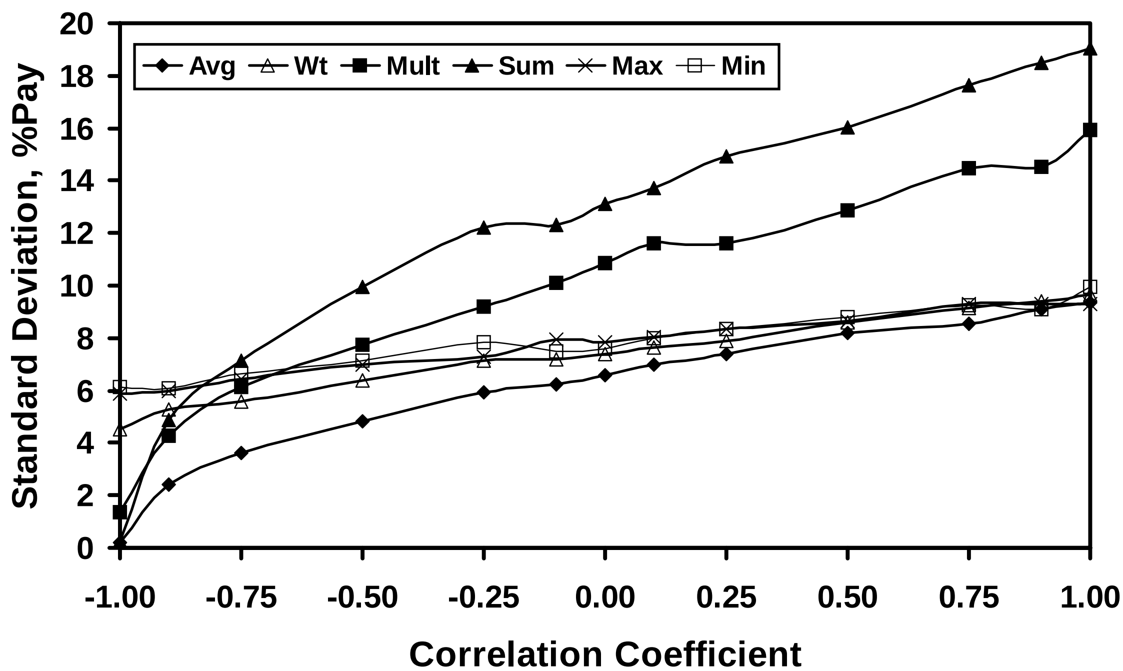


Figure 95b**. Comparison of simulation results for various methods for combining individual standard deviation payment factors for two populations with PWL = 50.**

CONCLUSION: TRADITIONAL COMPOSITE QUALITY MEASURES

The maximum and minimum methods essentially disregard the quality characteristics that do not give the maximum or minimum payment value, respectively. This seems to ignore useful information from the other quality characteristics. The averaging, multiplying, and summing methods for combining individual payment factors implicitly assume that each individual payment property is equally important. This is probably, but not always, a valid assumption.

Weighted average composite payment factors are intuitively appealing since it is very likely that all payment quality characteristics do not have an identical impact on pavement performance. A drawback to this approach, however, is that there is no obvious methodology for determining the appropriate weightings. The weightings, therefore, are subjective in nature and can vary from agency to agency depending on the agency or individual preferences.

Based on the analyses and discussion presented in this chapter, the weighted average method appears to be the best of the traditional methods for combining payment factors for individual quality characteristics into a single composite payment factor. However, as shown in the next chapter, there are other methods, based on predicted performance, for determining the final payment factor for a lot. The use of such methods is not widespread at this time.

11. RELATING PAYMENT TO PERFORMANCE

BACKGROUND

Decisions concerning payment relationships are extremely important since experience has shown that payment relationships in an acceptance plan are the most important factor from a contractor's perspective. The contractor submits a bid with a certain expectation of the amount of payment for the product. Achieving this amount of payment is critical in maintaining a viable business.

By the 1980s, many agencies had included payment reductions as part of the acceptance plans in their QA specifications. The intent was to penalize contractors for material or construction that did not fully comply with the specification requirements, but which was not of sufficiently poor quality to justify removal and replacement. Many agencies have now realized the need to also include incentive provisions (i.e., bonuses) to reward contractors for superior quality. These incentives and disincentive payment schedules were often based on engineering judgment, and similar test results could yield very different payment factors from one agency to another.

Relating quality and performance to payment is the most desirable form of payment relationship, because the relationship supports and defends the decision. This is true because negative payment adjustments are typically viewed with skepticism by the contracting industry. However, when the payment schedule can be shown to be related to quality and, preferably, to performance, it is viewed to be more credible than when it is established arbitrarily.

Two different approaches for developing performance models were considered during the current project. The first approach, which has been developed over many years through FHWA and National Cooperative Highway Research Program (NCHRP) support, is what has been called *performance-related specifications (PRS)*. These efforts have led to the development of PRS guide specifications and to computer programs that use theoretical and/or empirical models to predict pavement performance based on test results from the newly placed pavement. The software program PaveSpec, now in version 3.0, has been developed for PCC pavements and has been available to the general highway industry for a number of years. More recently, HMASpec has been developed for HMAC pavements.

The second approach that was considered is based on concepts similar to those of the PRS mentioned above. However, this approach is more general in that it can be applied to any set of quality characteristics for which a general relationship to pavement life can be developed or approximated. The second approach also considers only initial construction costs and major reconstruction or rehabilitation activities, and does not try to include routine or localized maintenance and repair costs into the payment determination.

Each of these is discussed briefly in the following sections. Much of the material on PRS is taken from unpublished manuscripts prepared by P.A. Kopac, T.M. Mitchell, and T.P. Teng of FHWA.

PERFORMANCE-RELATED SPECIFICATIONS (PRS)

An extensive research effort has been conducted to develop what are generally known as PRS. PRS attempt to use acceptance and payment plans that incorporate payment provisions based on pavement performance models. Current PRS are based on mathematical models that attempt to quantify pavement performance relationships among a number of important quality characteristics. The intent of these models is to provide a clearer understanding and better estimate of a pavement's performance than can be obtained by even the best intuitive engineering judgment. These performance models are then used to determine the payment adjustments in the acceptance plan.

The most elaborate PRS incorporate models to predict both pavement performance and maintenance costs. The performance prediction models use the pavement's design and materials information to predict when and to what extent the pavement will exhibit various types of distress, such as fatigue cracking or joint spalling. The maintenance cost models use predicted stress development to estimate post-construction life-cycle costs (LCC), which include the predicted costs of maintenance and rehabilitation that will be encountered throughout the projected life of the pavement.

As shown in figure 96, the inputs to the performance model include design variables such as traffic loading, climatic factors, drainage, and roadbed soil factors, and materials quality characteristics such as asphalt content, concrete strength, and pavement smoothness. These, in turn, are converted by the model into predicted occurrence and the extent of the various types of distress. This then becomes the input to the maintenance cost models whose output is ultimately a predicted LCC for the pavement.

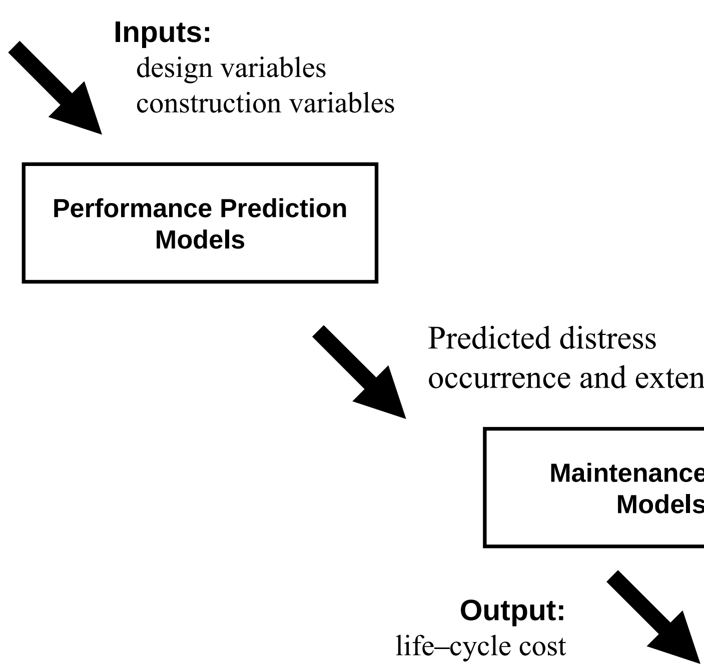


Figure 96. Flowchart of the PRS process (after Kopac, Mitchell, and Teng).

The target values for the quality characteristics are input to the models to determine the *as-designed LCC*. The actual measured values for the pavement's quality characteristics are then used as input to produce the estimated *as-constructed LCC*. The difference between the as-designed and as-constructed LCC values is then used as the basis for any incentive or disincentive payment.

Computer simulation routines are used with PRS to consider variability when determining payment adjustments. Mean and standard deviation values for the as-designed quality characteristics are entered into the simulation program along with other pavement design information. The values for as-designed mean and standard deviation can be determined from the design or based on agency policy or past experience. Hundreds, or preferably thousands, of iterations of the model are then performed.

For each iteration, the simulation routine selects from a normal distribution, defined by the as-designed mean and standard deviation values, for each designated quality characteristic. These simulated values are then input into the PRS models to arrive at an estimate for the as-constructed LCC for each iteration. The average of the many simulated LCC estimates is then used as the as-designed LCC.

The same process is then used with the mean and standard deviation from the acceptance sample test results to develop an estimated as-constructed LCC for each constructed lot. Payment factors are then based on the following equation:

$$PF = \frac{100\left[Bid + \left(LCC_{des} - LCC_{con}\right)\right]}{Bid} \tag{30}$$

where:
- PF = payment factor as a percentage of the bid price
- Bid = contractor's bid price
- LCC_{des} = as-designed LCC
- LCC_{con} = as-constructed LCC

Up to now, most PRS have been developed for individual projects. In that way, project-specific values for inputs, such as design, traffic, climate, etc., can be used. This project-by-project approach obviously provides the best estimate for pavement performance. An alternative would be to develop generic PRS for similar types or groups of projects, such as those within a given geographic or climatic region, or those with similar average daily traffic. While this would eliminate the need for developing a different PRS for each project, the predicted performance would not be as good an estimate as would a project-specific PRS.

To fully benefit from PRS, the performance models must be accurate and properly applied. The current models were developed based on generic national norms. Each agency that wishes to use the current PRS models must decide whether or not its local conditions deviate sufficiently from those norms to require revised performance models for its pavements.

PRS offer great promise to radically change the way in which payment plans are developed and payment factors are determined. However, they are, to a great extent, still in the pilot project stage and have not yet been widely adopted or employed by highway agencies. There are some potential roadblocks to their widespread adoption and implementation. One of these is that they can be perceived as being much more complex than any traditional specification with which an agency is familiar. This perception may be justified in light of the fact that 126 different inputs are possible when using PaveSpec 3.0. This may lead some agencies to be slow to adopt PRS.

While these perceptions of complexity may be overcome, another serious potential drawback is that PRS may require a significant amount of additional testing compared to current QA specifications. Regardless of how sophisticated the internal mathematical or empirical models in a PRS process are, the limiting factor for the accuracy and validity of the LCC prediction will still be the quality of the as-constructed information that is input. No matter how many iterations are simulated within a PRS model, the predictions will not be accurate if the as-constructed mean and standard deviation are not correct.

The large variability in the estimates, based on typical sample sizes = 3–5, has been shown in numerous places in preceding chapters. No matter how sophisticated and precise the models within the PRS, the limiting factor will be the size of the sample that the agency will take to determine the as-constructed means and standard deviations. If, because of personnel shortages, highway agencies are passing acceptance testing responsibilities to the contractors, it seems unlikely that agencies will be willing to increase the amount of acceptance testing that is performed.

Decision Regarding PRS

For a number of reasons, it was decided that this is not the right time to recommend PRS in a QA specifications development manual that is recommended for current use by all highway agencies around the country. Some of these reasons include:

- There is a wide array of input data that must be collected and maintained for input to the PRS.
- There is a possible perception that the PRS is a black box over which the agency has no control.
- Many agencies may not currently have the data or expertise to revise the generic models to fit the specific needs of their agencies.
- It seems unlikely that a large number of agencies would be willing to increase the amount of acceptance testing at this time.

It is believed that PRS are probably the specifications of the future (whether this is the immediate future or sometime later is still open for debate). While the promise of PRS was recognized, it was decided that a simpler approach for incorporating LCCs into payment determinations would be best for inclusion in the QA manual under development. The approach that was selected for the QA manual is presented in the next section.

LCC BASIS FOR PAYMENT ADJUSTMENTS

The concepts incorporated into the PRS discussed above are sound. The method that was selected for the QA manual is based on similar LCC concepts, but is easier to implement and should be easier to modify for a specific agency than is the PRS approach described above. The method considers only major repairs (such as resurfacing) and does not attempt to include routine or localized repair costs when calculating the LCC. The general approach can be applied to any agency, provided that there is some model or method for predicting pavement life before an overlay is required. Any existing model, including sophisticated PRS models, can be used to develop the predicted lives.

LCC Approach

Ordinarily, a pavement is designed to sustain a specified number of load applications before major repairs (such as resurfacing) are required. If, because of construction deficiencies, the pavement is not capable of withstanding the design load, it will fail prematurely. The need to repair this pavement at an earlier date results in an additional expense that must be borne by the agency since it usually occurs long after any contractual obligations have expired. Therefore, one possible purpose for an adjusted payment schedule might be to withhold sufficient payment at the time of construction to cover the extra costs anticipated in the future as a result of work that is of deficient quality.

Pavements are usually designed to withstand a required number of equivalent single-axle loads (ESALs). For those quality characteristics used in the design procedure, the as-built values can be compared to the design values to estimate the fraction of the design load that the pavement is capable of sustaining. As an approximate estimate, this fraction can be multiplied by the design life to obtain the expected life of the pavement. If greater precision is desired, a traffic growth rate can be assumed, the effect of which is to extend slightly the expected life (since fewer of the allowable loads will occur in the early part of a pavement's life).

To estimate the cost to the agency of premature pavement failure, it is necessary to determine the net present value of the various actions made necessary by early failure. For example, suppose that experience has shown that a new pavement typically lasts about 20 years and an overlay about 10 years. If the initial surface were to fail 1 or 2 years prematurely, it is not likely that an agency would do a minor repair to extend the life of the pavement to the originally expected value of 20 years. A much more practical decision would be to reschedule the overlay that was planned for the 20th year and do it 1 or 2 years sooner. However, if the overlay that was planned for the 20th year is rescheduled to an earlier date and overlays typically last 10 years, then all future overlays must be moved to an earlier time as well.

The procedure involves the calculation of a series of debits and credits and turns out to be relatively easy. Moving the overlay that was planned for the 20th year to the 18th year, for example, would result in a debit in net present value terms because it represents a cost in the 18th year that was not planned. However, there will also be a credit for no longer having to do an overlay in the 20th year. Since the overlay that was planned for the 20th year is now later in the future, the credit for this action is discounted to a greater degree, resulting in a net debit for the rescheduling of the overlay that was planned for the 20th year. This is illustrated in figure 97.

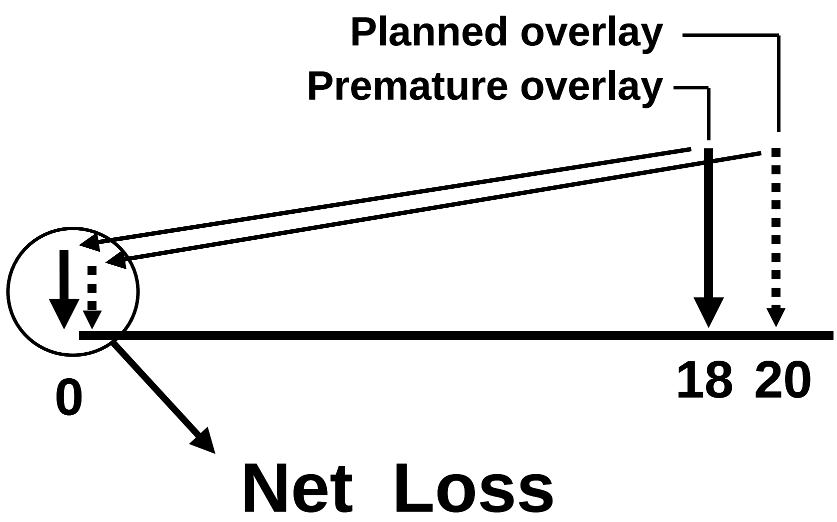


Figure 97. Illustration of the net impact of rescheduling an overlay 2 years earlier than originally planned.

While it is true that the net debits from the rescheduling of overlays farther in the future are discounted to a greater extent, and soon become insignificant, ignoring them altogether would substantially underestimate the true cost of pavement failure. Alternatively, selecting a specific analysis period would require an assumption about the residual value of a partially depleted overlay (information that is not readily available). Fortunately, this is an easy problem to solve mathematically to yield equation 31, the derivation of which is presented in the QA manual:[(1)]

$$PAYADJ = \frac{C(R^D - R^E)}{(1 - R^O)} \tag{31}$$

where:

	PAYADJ	=	appropriate payment adjustment for new pavement or overlay (same units of cost as C)
	C	=	present total cost of resurfacing
	D	=	design life of pavement or initial overlay
	E	=	expected life of pavement or overlay (independent variable)
	O	=	expected life of successive overlays
	R	=	(1 + *INF*)/(1 + *INT*)
	INF	=	long-term annual inflation rate in decimal form
	INT	=	long-term annual interest rate in decimal form

Performance Relationships

To apply the LCC basis for payment schedules in a manner that is both fair to all parties and legally defensible, it is necessary to have at least an approximate performance relationship. The purpose of the performance relationship is to predict from the quality characteristics measured at the jobsite what the expected service life of the construction item will be. This is the independent variable to be entered into the LCC equation presented in the previous section.

As noted in previous sections, efforts are proceeding on the development and implementation of extremely sophisticated computerized procedures to develop performance relationships and appropriate payment schedules. However, the successful completion, validation, and widespread adoption of these procedures are still some time away. Even when completed, the data requirements and the level of complexity of these procedures may deter their widespread use by practitioners seeking more practical methods that are easier to understand and apply.

Therefore, there is a need for an alternative approach for those agencies that choose to develop their own procedures, in their own way, that are tailored to their own specific circumstances. Perhaps more importantly, this alternative method needs to be sufficiently straightforward and scientifically sound so that agency engineers could not only understand it and use it with confidence, but could also modify it when necessary and be able to present it convincingly to the contractors whose work it will govern.

The derivation of such an alternative approach for developing simplified performance models is presented in depth in the QA manual along with a number of examples.[(1)] This derivation resulted in equation 32, which is a simple exponential model that is applicable to a wide variety of quality characteristics:

$$EXPLIF = Ae^{-(B_1\mathrm{PD}_1 + B_2\mathrm{PD}_2 + \cdots + B_k\mathrm{PD}_k)} \tag{32}$$

where:

	$EXPLIF$	=	expected life, in years
	A	=	constant to be determined
	B_k	=	coefficients to be determined for each of the k quality characteristics
	PD_i	=	percent defective of individual quality characteristics
	k	=	number of quality characteristics
	e	=	base of natural logarithms

This model has certain important advantages. It tends to produce an “S” shape that is believed to be an appropriate form for many performance relationships. Also, because this particular model produces a maximum of *A* and a minimum as close to zero as desired (but not below zero), it can easily be made to fit most real-world situations. Finally, it requires relatively straightforward data and simple mathematics to accommodate as many acceptance characteristics as are likely to be necessary.

For example, consider a resurfacing project for which historical data have shown the typical expected life to be about 10 years. A typical value for the AQL is PD = 10, while RQL values tend to vary more widely, depending on what quality level the agency believes justifies removal and replacement at the contractor’s expense. For the purposes of this example, suppose that the agency has decided to use RQL values of PD = 65, 75, and 85 because it is believed that these levels correspond to approximately a 50-percent loss in pavement life (or an expected life of 5 years). These assumptions lead to the completed data matrix shown in table 55.

Table 55. Completed data matrix for the example of an exponential model.

PD_{VOIDS}	PD_{THICK}	PD_{SMOOTH}	*EXPLIF* (years)
10 (AQL)	10 (AQL)	10 (AQL)	10 (AQL)
65 (RQL)	10 (AQL)	10 (AQL)	5 (poor voids)
10 (AQL)	75 (RQL)	10 (AQL)	5 (poor thickness)
10 (AQL)	10 (AQL)	85 (RQL)	5 (poor smoothness)

All that remains is to use the information in the data matrix to solve for the unknown coefficients in equation 32 (the exponential performance equation for *EXPLIF*). To accomplish this, it is first necessary to take logarithms of both sides, producing equation 33:

$$ln(EXPLIF) = ln(A) - B_1 PD_{VOIDS} - B_2 PD_{THICK} - B_3 PD_{SMOOTH} \tag{33}$$

where:
- $EXPLIF$ = expected life, in years
- PD_{VOIDS} = air voids percent defective
- PD_{THICK} = thickness percent defective
- PD_{SMOOTH} = smoothness percent defective
- A, B_1, B_2, B_3 = unknown coefficients
- ln = natural logarithm operator

The values from each row in table 55 can be inserted into equation 33 to develop four equations that can be used to solve for the four unknown coefficients, leading to the following performance model:

$$EXPLIF = 13.8e^{-(0.0126\mathrm{PD}_{VOIDS} + 0.0107\mathrm{PD}_{THICK} + 0.00924\mathrm{PD}_{SMOOTH})} \tag{34}$$

This model would then be used to develop the expected-life value to be used in equation 31 to determine the payment adjustment, which could be either positive or negative depending on the expected-life estimate.

The concepts used in the development of equation 34 have widespread application. The only requirement is that the agency creates a data table such as table 55. Any method with which the agency is comfortable can be used to develop the values for estimated pavement life that result from the various levels of the quality measure. For example, if a performance model is available and the highway agency has confidence in the predictive capability of the model, then it could be used to develop the estimated expected life of the pavement.

Composite Quality Measure

As noted in chapter 10, specifications based on multiple quality characteristics frequently use payment equations that include a separate term for each of the quality characteristics so that the resultant payment adjustment is a function of the combined effects of all of the quality measures. An alternate method to accomplish the same purpose is to base the payment equation on a single quality measure that is a composite of the individual quality measures. This latter approach, because it keys the various decisionmaking steps to a single performance indicator, simplifies the procedure and offers several practical advantages.

The benefit of the composite quality measure is shown in a simple example that uses two acceptance quality characteristics—air voids and thickness. Suppose that the expected-life model in equation 35 has been developed for these two characteristics:

$$EXPLIF = 22.9 - 0.163\ PD_{VOIDS} - 0.135\ PD_{THICK} + 0.000961\ PD_{VOIDS} \times PD_{THICK} \quad (35)$$

Since the quality measure in equation 35 is PD, which ranges from a minimum of zero to a maximum of 100, it will be convenient to develop a composite quality measure (*PD**) that spans that same range. As derived, the value of *EXPLIF* in equation 35 ranges from 2.71 to 22.9. By algebraic operations, equation 35 can be modified to span the range from 0 to 100, thereby yielding equation 36:

$$PD^* = 0.807\ PD_{VOIDS} + 0.669\ PD_{THICK} - 0.00476\ PD_{VOIDS} \times PD_{THICK} \quad (36)$$

where:
- PD^* = composite quality measure, in units of percent defective
- PD_{VOIDS} = air voids percent defective
- PD_{THICK} = thickness percent defective

PD^* progresses smoothly from 0 to 100 percent as the individual quality measures (PD_{VOIDS} and PD_{THICK}) vary throughout the same range. Table 56 presents a few selected examples of this:

Table 56. Examples of computed *PD values for selected individual PD values.**

PD_{VOIDS}	PD_{THICK}	*EXPLIF* (years)	PD^*
0	0	22.9	0.0
10	10	20.0	14.3
50	50	10.4	61.9
25	75	10.5	61.4
100	100	2.7	100.0

As shown in table 56, the case in which PD_{VOIDS} and PD_{THICK} are both equal to 50 produces essentially the same level of expected life as the case in which $PD_{VOIDS} = 25$ and $PD_{THICK} = 75$. This result flows directly from the manner in which the *EXPLIF* equation was derived and is realistic because an increase in the quality of one measure might be expected to offset a decrease in the quality of the other. Appropriately, both cases produce virtually the same value of *PD** in the last column of the table, indicating that *PD** is well suited as a measure on which a QA specification can be based.

This property of the composite quality measure, which properly accounts for the combined effects of multiple quality characteristics, also makes it possible to develop an RQL provision that is far superior to the alternative of defining separate RQL provisions for the individual quality measures. For the example in table 57, it is assumed that the agency has defined for air voids and thickness separate RQL provisions of $PD_{VOIDS} \geq 75$ and $PD_{THICK} \geq 90$. Clearly, case 3 in table 57 is, by far, the worst case, yet it is not recognized as an RQL condition when using individual RQL provisions, while the other two cases are considered as such.

Table 57. Illustration of the problem with separate RQL provisions.

Case	Quality Level		Reject?	*PD**
	Air Voids	Thickness		
1	PD = 75 (RQL)	PD = 0 (excellent)	Yes	60.5
2	PD = 0 (excellent)	PD = 90 (RQL)	Yes	60.2
3	PD = 74 (almost RQL)	PD = 89 (almost RQL)	No	87.9

To demonstrate the effectiveness of an RQL provision based on the composite quality measure, equation 36 was used to compute the corresponding values for PD^* that appear in the last column of table 57. In this example, a PD^* of 60 or more would be regarded as rejectable and, as shown in the last column, case 3 is properly recognized as being well into the rejectable region.

All of the topics covered in this section are covered in much greater detail and with more examples in the QA manual.[(1)] This information is included in this report merely to show the relative simplicity of this method and to highlight some of its benefits.

12. BRIEF SUMMARY AND RECOMMENDATIONS

PROJECT OBJECTIVE

The objective of the project was to develop a comprehensive QA manual, supported by scientific evidence and statistical theory that provides step-by-step procedures and instructions for developing effective and efficient QA specifications. This technical report summarizes what steps were taken to accomplish this goal and the analyses that were conducted to support the recommendations made in the QA manual.

While the focus and objectives of the two documents are quite different, they are not totally stand-alone documents. As such, the technical report should be read in conjunction with and as a companion to the QA specifications manual that also resulted from the project.

QA SPECIFICATIONS DEVELOPMENT PROCESS

One of the major accomplishments of the project was the development of the three-phase QA specifications development process that is presented in chapter 3. The process was presented to and approved by the panel of representatives from the pooled fund States that provided the funding for this project. The process is presented in flowchart form for ease of understanding.

RESULTS OF DETAILED ANALYSES

After seeing and approving the specifications development process flowcharts, the panel was surveyed to identify those topics that they thought were most important for further detailed analyses. The items selected by the panel also comprise some of the major decisions that are necessary during the specifications development process. These items are presented below, along with the recommendations that were included in the QA manual:(1)

What quality measure should be used for individual quality characteristics?

From the many potential quality measures that were identified, three were chosen for detailed analyses—PWL (and its complement PD), AAD, and CI. After initial analyses, CI was eliminated because it offered no benefits over AAD and because it provided a slightly biased estimate for the population CI. Further studies showed that PWL and AAD generally performed comparably to one another. PWL was recommended because it provides a direct measure of variability, whereas AAD provides, at best, an indirect measure of variability. In fact, the same population AAD could apply to very different populations. PWL also works effectively with both one-side and two-sided specifications, while AAD only applies to two-sided specifications that have a definite target value.

What payment relationships should be used for individual quality characteristics?

The selection of the payment equation is, to some extent, a subjective decision for each agency based on the level of risk that it considers reasonable. A general method was developed to determine the equivalent PWL and AAD values for any given value of PWL or AAD. From this it was possible to show that separate PWL and AAD payment equations could be developed to provide equivalent EPs for the same population. While no single payment relationship was recommended for individual quality characteristics, the extensive discussion in this report regarding calculating and evaluating risks provides the agency with the tools and techniques to develop payment equations that they believe to be fair and reasonable.

How should multiple quality characteristics be combined into a single payment factor?

Analyses were conducted on six different methods for combining payment factors for individual quality characteristics into one combined payment factor:

- Averaging the individual payment factors.
- Using a weighted average of the individual payment factors.
- Multiplying the individual payment factors.
- Summing the individual payment adjustments.
- Using the maximum individual payment factor.
- Using the minimum individual payment factor.

The analyses showed that the variability of the combined payment factors for two characteristics was related to whether they were positively correlated, negatively correlated, or not correlated. For four of the methods, the EP values did not vary with the correlation between the variables. However, the EP did vary with the correlation when using either the maximum or minimum individual payment factor.

Another method—using a single composite quality measure derived from a general performance model to predict expected pavement life—was developed and is the recommended approach. If an agency wishes to use a simpler method for determining the composite payment factor, the first four methods listed above are candidates. One of the two averaging methods might be preferred since the averaging process has the effect of reducing the variability in the estimated payment factors.

What procedures should be used to verify the contractor's test results if they are to be used in the acceptance and payment decision?

A detailed discussion and analysis were presented regarding the risks involved in the various approaches to verifying the contractor's test results. OC or power curves, surfaces, and tables were presented to illustrate the risks associated with the different procedures and sample sizes. The relatively high risks that are associated with typical agency verification testing frequencies were pointed out.

A discussion was presented concerning which verification procedures were appropriate for use with split and independent samples. For split samples, the recommended approach is to use a maximum allowable limit for the difference between the agency's results and the contractor's results for each individual pair of split samples, but also to accumulate the results for use in a *t*-test for paired measurements. The best approach for independent samples is to use the *F*-test to compare the variances in the contractor's tests and the agency's tests. The result of the *F*-test then determines the manner in which the *t*-test is conducted to compare the means of the contractor’s tests and the agency’s tests. Each agency can select its own testing frequency based on its assessment of the risks provided in the OC curves that were developed during this project.

APPENDIX A: ANNOTATED BIBLIOGRAPHY OF SELECTED ITEMS FROM THE INITIAL LITERATURE SEARCH

Afferton, K.C.; Freidenrich, J.; and Weed, R.M. "Managing Quality: Time for a National Policy," *Transportation Research Record 1340*, 1992, pp. 3-39.

Statistical quality assurance (SQA) (currently in use or under development in approximately three-quarters of the States) has proven to be a very effective tool for encouraging high-quality construction. However, although statistical specifications writing must now be recognized as a thoroughly scientific activity, there is great disparity in its application on the State level and many current practices and published standards are far from optimal. Part 1 of this paper stresses the need for sweeping reforms and suggests that the establishment of a uniform national policy on transportation QA is overdue. Part 2 describes a variety of obstacles that must be overcome if such a transformation is to be made. Part 3 outlines an extensive series of principles that must be understood in order to derive the maximum benefit from a QA program. And finally, Part 4 presents a plan of action that, if conscientiously followed, will significantly increase the effectiveness of transportation QA practices nationwide.

Allen, O.B. and Newcombe, P.A. "A Three-Class Procedure for Acceptance Sampling by Variables," *Technometrics*, Vol. 30, No. 4, 1988, pp. 415-421.

A three-class procedure for acceptance sampling by variables is introduced as an alternative to both the three-class attributes plan and the two-class variables plan. The procedure, which requires that the quality characteristic be normally distributed, has the advantage of requiring a smaller sample size than a three-class attributes plan with approximately the same OC surface. The advantage of the three-class variables plan over the two-class variables plan is the ability to discriminate between lots with high and low percentages of marginally conforming products. Two equivalent methods of stating the decision rules for the plan are suggested. It is shown that the OC surface for the three-class variables procedure may be constructed with a special case of bivariate, noncentral *t*-distribution. The motivation and methods for choosing a plan are discussed and illustrated.

Aminzadeh, M.S. "Inverse-Gaussian Acceptance Sampling Plans by Variables," *Communications in Statistics: Theory and Method*, Vol. 125, No. 5, 1996, pp. 923-935.

Variable sampling plans to control the defective fraction are obtained using the Inverse-Gaussian distribution. OC curves are obtained and the impact of the sample size and the specification limits on these curves are discussed. Simulation studies are used to investigate the sensitivity of the sampling plans under the more commonly used normal distribution.

Amirkhanian, S.N.; Burati, J.L.; and Mirchandani, H.C. "Effect of Testing Variability on Contractor Payment for Asphalt Pavements," *Journal of Construction Engineering and Management*, Vol. 120, No. 3, September 1994, pp. 579-592.

Acceptance of asphalt pavements on a lot-by-lot basis depends on how well the material conforms to specified quality characteristics. If a lot does not fall within the specified tolerance limits for a given characteristic, the payment to the contractor for that lot may be adjusted according to a predetermined payment schedule. This study evaluated the effects of the variability of the test method on the payment to the contractor. Payment factors based on selected State highway agency specifications, typical of the most commonly used type of acceptance plans, were estimated using simulation techniques. The results were illustrated by plotting the OC in terms of the expected payment. The OC plots show that the contractor might receive a reduced payment even when there is no material variability and the process mean is on target. This reduction was a result of the relatively large magnitude of the testing variability as compared with the tolerance range specified in the simulated acceptance plans. With suitable modifications, the simulation routine developed to estimate contractor payment can be used for any acceptance plan.

Balamurali, S. and Kalyanasundaram, M. “Determination of an Attribute Single Sampling Scheme,” *Journal of Applied Statistics*, Vol. 24, No. 6, 1997, pp. 689-695.

A procedure for the selection of a new sampling scheme, called the *Single Sampling Scheme*, is presented. Additionally, a table for the selection of a single sampling scheme, which is indexed by various combinations of entry parameters, is given. The method of table construction and the advantages of single sampling schemes are discussed.

Barros, R.T.; Weed, R.M.; and Willenbrock, J.H. “Software Package for Design and Analysis of Acceptance Procedures Based on Percent Defective,” *Transportation Research Record 924*, 1983, pp. 85-93.

The trend toward statistical end-result specifications (ERS) has led to the development of construction specifications based on the concept of PD. To analyze the risks and determine the effectiveness of the acceptance procedures associated with these specifications, OC curves must be constructed. However, many potential users do not have a working knowledge of the noncentral *t*- and beta distributions necessary for this development. The underlying theory, several useful references, and a conversational computer program that greatly simplifies the design and analysis of specifications of this type are presented.

Benson, P.E. “Comparison of End Result and Method Specifications for Managing Quality,” *Transportation Research Record 1491*, 1995, pp. 3-10.

The results of a statistically designed experiment in which AC cores and nuclear gauge readings were taken from five California projects are reported. Relative compaction for the projects was controlled with a method specification. Analysis of variance is used to separate test error and the locational components of variance for specific gravity, asphalt content, lift thickness, and grading. Compaction results are compared with similar results from 16 ERS jobs studied previously. Relative compaction on the end-result jobs averages 3.1 percent higher in value. Findings on test precision, increased lot size, and material variability are discussed.

Benson, P.E. “Performance Review of a Quality Control/Quality Assurance Specification for Asphalt Concrete,” *Transportation Research Record 1654*, 1999, pp. 88-94.

A statistical review of 50 jobs recently completed by using California’s QC/QA specification for AC is presented. Performance is contrasted to the quality achieved under method specification and ERS. A cost analysis is made and issues related to verification are discussed. Improvements to the current specification are proposed. The data present clear evidence that the allowable tolerance of ±0.5 percent for asphalt content is too broad for current practices. Also, an increase in compaction variability for many QC/QA jobs could be controlled by adding an upper specification limit or adopting a two-sided volumetric specification. The cost for QC/QA jobs went up approximately 3 percent to pay for bonuses allowed under the specification. An analysis of the contractor’s QC test data indicate that this increase is more than compensated for by projected reductions in future rehabilitation costs. However, a significant lack of agreement between the contractor QC and agency QA tests brings this finding into question. More rigorous verification of contractor-provided test results must be incorporated into the specification and the results must be analyzed before cost-effectiveness can be determined.

Budwig, J.L. “Bituminous Pavement Smoothness: Statistically Based Approach to Acceptance,” *Transportation Research Record 1544*, 1996, pp. 125-134.

Since 1987, the Federal Lands Highway (FLH) branch of FHWA has been evaluating acceptance of newly constructed bituminous pavements using California-type profilograph measurements. California Test Method 526 and FLH T504, as well as various acceptance plans, have been used in this evaluation. This study examined: (1) whether operator trace reduction variability was too large for acceptance testing, (2) which type of acceptance plan should be incorporated into the *Standard Specifications for Construction of Roads and Bridges on Federal Highway Projects* (FP 92), and (3) two commercially available computerized trace reduction systems. The study concludes that when used in conjunction with statistical evaluation procedures, the test method is suitable for acceptance purposes and that computerized trace reduction is superior to manual reduction. Also presented are some fundamentals of statistically based acceptance that are not widely known or understood by highway engineers.

Burati, J.L.; Antle, C.E.; and Willenbrock, J.H. "Development of a Bayesian Acceptance Approach for Bituminous Pavements," *Transportation Research Record 924*, 1983, pp. 64-71.

Traditional approaches for estimating the percentage of a lot that is within specification limits (PWL) are based on random samples taken from the lot being evaluated. These approaches suffer from the small sample size necessitated by the destructive and time-consuming tests that are usually used in determining the quality of the materials. The development of a Bayesian approach for estimating PWL, which incorporates information concerning the contractor's past performance on the project and the current sample results that determine the estimate for the PWL of the current lot, is presented. The procedure assumes that the daily population mean is a random variable that follows a normal distribution, that the production process is also normally distributed, and that the process variance is constant. These assumptions are confirmed by using goodness-of-fit tests on data collected from 13 bituminous runway paving projects. Computer simulation shows that the Bayesian PWL estimators are slightly biased in comparison with the traditional quality index method, but that the PWL estimators exhibit smaller variances than the traditional method.

Burati, J.L.; Bridges, W.C.; and Ackerman S.A. "Evaluation of Quality Assurance Programs for Bituminous Paving Mixtures," Report No. FHWA-SC-95-002, South Carolina DOT, May 1995.

The South Carolina DOT QA program currently consists of the Record Sample, Independent Assurance (known as the *Green Tag* program), and Split-Sample programs. Because these three programs were developed at different times and possibly for different reasons, a study was needed to determine how the three programs meshed together, if there are overlaps or gaps in these programs, and if they are serving the function or functions for which they were intended. The study consisted of: (1) a review of the documentation of the existing programs, (2) interviews with FHWA and South Carolina DOT personnel to establish the objectives for each program, (3) a survey of other State highway agency independent assurance practices, and (4) an analysis of historical data from each of the programs. ASTM precision statements were also considered when evaluating the allowable tolerances for comparing test results. This report recommends that the Record Sample and Split-Sample programs be maintained with some modification in the procedures, including using independently obtained samples (as opposed to split samples) for the Record Sample program and slightly modified allowable comparison tolerances for both programs. It was concluded that the Green Tag program is not necessary for the purpose of independent assurance. It primarily serves as an enforcement function and is not needed unless the South Carolina DOT believes that it is necessary to provide an external psychological inducement to ensure that its inspectors properly perform their job functions.

Burati, J.L. and Willenbrock, J.H. "Acceptance Criteria for Bituminous Surface Course on Civil Airport Pavements," Report No. FAA-RD-79-089, Federal Aviation Administration, 1981.

Research was undertaken to extend the use of statistically based airport pavement materials specifications that incorporate price-adjustment features. During the course of the project, data on the physical characteristics of the pavement materials were collected from 13 airport pavement construction projects. A statistical analysis of this data permitted the determination of the parameters (mean and standard deviation) of existing airport projects, and these parameters were then used to develop acceptance plans and price-adjustment factors. OC curves and the curves of expected payment were used to determine the appropriate acceptance plans, which were based on the percentage of material falling within the specification limits (PWL). By using a continuous rather than a discrete price-adjustment schedule, it was possible to avoid the problem of large differences in payment associated with relatively small differences in quality (as measured by the PWL). A computer program was developed to approximate the expected payment curves associated with different continuous price-adjustment systems. This program is applicable to one-sided specification limits such as density. For properties such as air voids, which require both an upper and lower specification limit, the OC curves were determined by computer simulation of 10,000 randomly drawn samples.

Chang, L-M. and Hsie, M. "Developing Acceptance Sampling Methods for Quality Construction," *Journal of Construction Engineering and Management*, Vol. 121, No. 2, 1995, pp. 246-253.

The role that the acceptance sampling method plays in designing the QA specification is discussed. The acceptance sampling method applies statistics to specify the number of measurements needed and determines how to make an acceptance or rejection decision based on measured data. Four acceptance sampling methods are presented, including the variable single sampling method, quality index sampling method, attribute double sampling method, and attribute proportion sampling method. The theories are derived and the applications of the four quality acceptance sampling methods are described. The advantages and disadvantages of the different sampling methods are compared.

Darter, M.I.; Hoerner, T.E.; Smith, K.D.; Okamoto, P.A.; and Kopac, P.A. “Development of Prototype Performance-Related Specification for Concrete Pavements,” *Transportation Research Record 1544*, 1996, pp. 109-115.

The development of a prototype PRS for concrete pavement is summarized. The prototype PRS requires that the pavement lot be divided into consistent sublots for the measurement of quality characteristics that are then used to estimate future performance and life-cycle costs. The difference between the life-cycle costs of the target (as-designed) pavement and the actual (as-constructed) pavement lot is used to determine a rational pay adjustment. Both the means and the variations of all of the quality characteristics are directly considered in the pay factor determination. A Microsoft® Windows®-based computer program, PaveSpec, was developed for use with the specification simulation and for generating pay adjustments. However, additional work is required to make this a fully practical PRS.

Dobrowolski, J. and Bressette, T. “Development of Quality Control/Quality Assurance Specifications by Using Statistical Quality Assurance for Asphalt Concrete Pavements in California,” *Transportation Research Record 1632*, 1998, pp. 13-21.

In 1996, Caltrans implemented QC/QA specifications for AC paving. These specifications require contractor QC and provide rewards or penalties based on statistical quality analysis of eight quality characteristics. These specifications were developed through a joint Caltrans/industry group and are supplemented with a QC manual. In March 1996, the first project using the specifications went to bid and, in the first year, six projects were completed. Since then, revisions to the specifications have been developed. The specifications and manual issues, recommendations based on 1996 projects, changes that Caltrans has made, and anticipated additional endeavors in areas of QC/QA and materials testing are discussed.

Douglas, K.D.; Coplantz, J.; Lehman, R.; and Bressette, T. “Evaluation of Quality Control/Quality Assurance Implementation for Asphalt Concrete Specifications in California,” *Transportation Research Record 1654*, 1999, pp. 95-101.

Caltrans has implemented a new QC/QA specification for AC pavement. As part of this implementation, Caltrans realized the need for objective feedback on its design and implementation, and selected Nichols Consulting Engineers to perform this study. The study objective, the research approach used, a description of the projects included in the study, and recommendations and findings are presented. Lessons learned from the Caltrans experience are shared to the benefit of other States implementing QC/QA specifications for AC pavement.

Elliott, R.P. “A Value Concept for Pavement Construction Pay Adjustment Schedules,” *Transportation Research Record 1040*, 1985, pp. 45-48.

A value concept is presented that can serve as a basis for developing rational payment schedules for pavement construction. Provisions are made for incorporating both the average and the standard deviation of the materials tests into a payment determination scheme that is based on the relative pavement life effects. The concept is based on the recognition that, at the time the pavement is considered to have failed, only a small percentage of the surface actually exhibits severe distress. As a result, the life of the pavement is controlled not by the average or 50th percentile of the material, but by a lower percentile representative of the actual surface distress.

Elliott, R.P. and Qiu, Yanjun. “Analysis of Contractor Pay Adjustment Schedule Using Simulation,” *Transportation Research Record 1544*, 1996, pp. 109-115.

A common provision in QC/QA construction contracts is the adjustment of the contractor’s pay on the basis of the quality of the construction. The expected impact of the provision on payment should be examined to ensure that the adjustments are neither unduly severe nor excessively lenient. Most pay adjustment plans have been

developed around a quality index by using a PD approach. Analyses of these plans are complex, but reasonably well defined. Other plans are more complex and do not lend themselves to direct analysis. Computer simulation can be used to examine these plans. An example is given. It is shown that the simulation process can provide a better, more detailed examination of the pay schedule than is possible by simply determining the expected payment. In particular, the simulation process can indicate the variability of payment at various quality levels and can identify the factors most responsible for the pay adjustment.

Fan, D-Y. "Bayesian Acceptance Sampling Scheme for Pass-Fail Components," *Communications in Statistics: Theory and Method***,** Vol. 20, No. 8, 1995, pp. 2351-2355.

An approach is given for determining the choice of a prior distribution for the design of a Bayesian acceptance sampling scheme. The appropriate prior distribution is selected from confidence coefficients corresponding to classical lower confidence bounds. A numerical example is provided.

Gentry, C. and Yrjanson, W.A. "Specifications for Quality Control: A Case Study," *Transportation Research Record 1126*, 1987, pp. 37-41.

One of the essential qualities of a specification is reasonableness. Court decisions and their economic consequences demand that specifications be based on reasonable requirements. Specifications that call for unnecessary perfection are hardly reasonable; furthermore, they do not ensure performance. Specifications that attempt to control quality through extremely limited tolerances may, in fact, be counterproductive. When QC efforts are directed to comply with such specifications, quality may be compromised, contract administration may be difficult, and additional costs may be incurred, all without improving the performance of the completed work. A case study is presented to illustrate the problems created by an excessively restrictive specification. Alternatives and comments for improvement of the specification are offered.

Gibra, I.N. "Recent Developments in Control Chart Techniques," *Journal of Quality Technology*, Vol. 7, No. 4, 1975, pp. 183-192.

In the last three decades, several control chart procedures were developed. Prominent among these are the Cumulative Sum, Economic Design of X-Control, and Acceptance and Multicharacteristic charts. Descriptions and analyses of these charts are presented in an attempt to bring past research into current perspective.

Govindaraju, K. "Certain Observations on Lot-Sensitive Sampling Plan," *Communications in Statistics: Theory and Method*, Vol. 19, No. 2, 1990, pp. 617-627.

A comparison of the sample-size efficiency of the lot-sensitive plan with the double and multiple sampling plans is presented. It is shown that a fully curtailed lot-sensitive plan will involve a smaller average sample number than the equivalent double and multiple sampling plans.

Govindaraju, K. and Kuralmani, V. "A Note on the Operating Characteristic Curve of the Known Sigma Sampling Variables Plan," *Communications in Statistics: Theory and Method***,** Vol. 21, No. 8, 1992, pp. 2339-2347.

The OC curves of certain known sigma sampling plans may not be satisfactory in that they have a tendency to reject lots of acceptable quality. The theory and a method to identify such known sigma variables plans possessing unsatisfactory OC curves are presented.

Govindaraju, K. and Subramani, K. "Selection of Double Sampling Attributes Plan for Given Acceptable Quality Level and Limiting Quality Level," *Communications in Statistics: Simulation and Computation*, Vol. 21, No. 1, 1992, pp. 221-242.

Tables and procedures are given for finding the double sampling plan, conditional double sampling plan, link sampling plan, and ChSP-4 and ChSP-4A chain sampling plans involving a minimum sum of the producer's and consumer's risks for the specified acceptable quality level and the limiting quality level.

Govindaraju, K. and Subramani, K. “Selection of Single Sampling Attributes Plan for Given Acceptable Quality Level and Limiting Quality Level Involving Minimum Risk,” *Communications in Statistics: Simulation and Computation*, Vol. 19, No. 4, 1990, pp. 1293-1302.

A table and a procedure are given for finding the single sampling attributes plan involving a minimum sum of the producer’s and the consumer’s risk for the specified acceptable quality level and the limiting quality level.

Hamaker, H.C. “Acceptance Sampling for Percent Defective by Variable and by Attributes,” *Journal of Quality Technology*, Vol. 11, No. 3, 1979, pp. 139-148.

Various methods for adjusting variables and attributes sampling plans so that they possess nearly identical OC curves are reviewed and extended. It is demonstrated that the OC curve for the *S*-method plans can be adequately derived from a normal approximation and that the more complicated use of the noncentral *t*-distribution can be avoided. The relative efficiencies of the different types of single sampling plans are shown to be practically functions of the indifference quality alone. The relationships between these efficiencies and the choice of the specification limit are discussed in detail.

Hawkes, C.J. “Curves for Sample Size Determination in Lot-Sensitive Sampling Plans,” *Journal of Quality Technology*, Vol. 11, No. 4, 1979, pp. 205-210.

Curves are given for determining the required sample size for zero acceptance number sampling plans for specified values of lot size and lot tolerance PD. These allow a quick and easy assessment of the minimum required sample size to obtain lot PD protection for 1-percent, 5-percent, 10-percent, 20-percent, and 50-percent consumer risk.

Hughes, C.S. “Incentive and Disincentive Specification for Asphalt Concrete Density,” *Transportation Research Record 986*, 1984, pp. 38-42.

The background for a specification that includes both positive and negative price adjustments for the density of AC is presented. The results that have been obtained since the specification was introduced in 1978 are described. The incentive features of the specification are emphasized because it is believed that they are unique and have been the primary reason that improved densities have been obtained in Virginia for the past 6 years.

Hughes, C.S. “Variability in Highway Pavement Construction,” *NCHRP Synthesis 232*, 1996.

This synthesis addresses variability questions and the importance of defining the variability of materials and construction processes. For potential users of these measures of variability, an update of the typical materials and construction variability and the use of incentive and disincentive pay schedules for acceptance are presented.

Irick, P.E. “A Conceptual Framework for the Development of a Performance-Related Materials and Construction Specification,” *Transportation Research Record 1126*, 1987, pp. 1-27.

Pavement design and performance concepts that provide a systematic basis for the development of specifications for materials and construction (M&C) are presented. It is assumed that the conceptual framework for specifications includes eight sets of relationships among the process variables and nine sets of inputs or outputs for the relationships. Independent variables are selected that have predictable effects on performance-related output variables. From these independent variables, variables appearing explicitly in prediction functions (EPF) are selected and subdivided into traffic factors, environmental factors, and pavement structure factors. EPF variables can be replaced by surrogate variables (SPF) when M&C control for these secondary variables is easier to provide. Other secondary variables are the control factors (CF), which have predictable effects on the EPF of SPF variables. EPF variables related to the M&C process are denoted as MCF. In general, a stochastic prediction model consists of a prediction function that may be completely known from mechanistic considerations; may be partially known except for undetermined constants; or may be assumed to be a linear combination of linear, curvilinear, and interaction effects among independent variables. General forms of prediction equations for stress and distress, stress-load equivalence relationships, traffic prediction relationships, relationships among M&C specification factors, and performance-cost relationships are presented. Pavement design criteria and M&C specification factors are added as the initial conditions for the definition of a pavement design for a given requirement.

Kandahl, P.S.; Cominsky, R.J.; Maurer, D.; and Motter, J.B. "Development and Implementation of Statistically Based End-Result Specifications for Hot-Mix Asphalt in Pennsylvania," *Transportation Research Record 1389*, 1993, pp. 9-16.

In the past, Pennsylvania DOT used the concept of single samples and tests to determine the quality of hot-mix asphalt (HMA) mixtures. This study develops a statistically based ERS for HMA pavements that makes the contractor responsible for QC and Pennsylvania DOT responsible for QA. Field data from several HMA paving projects were analyzed statistically to establish realistic numerical limits for the various test parameters used in the specification. Three pay items (asphalt content, percent passing #200, and mat density) are included in the specification. The specifications have provided Pennsylvania DOT with a means of evaluating and comparing the dollar value of the year-by-year improvement in HMA quality.

Kirkpatrick, R.L. "Confidence Limits on a Percent Defective Characterized by Two Specification Limits," *Journal of Quality Technology*, Vol. 2, No. 3, 1970, pp. 150-155.

A quality problem is to assess the performance of a production process with respect to specifications. A solution is obtained by estimating PD and placing confidence limits on the PD for each product characteristic. Tables are presented for determining a point estimate of PD and 90-, 95-, and 99-percent confidence limits on the PD of a characteristic with both upper and lower or single specification limits. On the assumption of normality, only the specification limits, sample size, sample mean, and standard deviation are needed to give a solution to the problem.

Kopac, P.A. "Current Practices in Acceptance of Bituminous Concrete Compaction," *Transportation Research Record 986*, 1984, pp. 43-46.

Current procedures employed by State highway agencies to determine the acceptability of bituminous concrete compaction are discussed. Both statistical and nonstatistical acceptance plans are covered. Statistical acceptance plans can provide a clear indication of the quality levels that are desired with the estimation of the construction quality. However, many of the statistical acceptance plans currently in use are inadequate because they are inefficient or statistically unsound, or both. Also, there is a considerable lack of uniformity among acceptance plans and the price adjustment schedules they contain. Recommendations are made to improve acceptance plans.

Kuralmani, V. and Govindaraju, K. "Modified Tables for the Selection of the Double Sampling Attributes Plan Indexed by AQL and LQL," *Communications in Statistics: Theory and Method*, Vol. 24, No. 7, 1995, pp. 1897-1921.

Modified tables are presented for the selection of double sampling plans for a given AQL, producer's risk, LQL, and consumer's risk, giving the minimum sum of the average sample numbers at the AQL and LQL under the conditions of the Poisson model for the OC curve.

Kuralmani, V. and Govindaraju, K. "Selection of Conditional Sampling Plans for Given AQL and LQL," *Journal of Applied Statistics*, Vol. 20, No. 4, 1993, pp. 467-479.

Procedures and tables for the selection of conditional sampling plans are given. The sampling plans discussed include the conditional double sampling plan, chain-deferred sampling plan, link sampling plan, and ChSP-4A chain sampling plan. The procedures and tables for each are provided for a given AQL, producer's risk, limiting quality level (LQL), and consumer's risk, giving a minimum sample size at AQL under the conditions of the Poisson model for the OC curve.

Moore, R.M.; Mahoney, J.P.; Hicks, R.G.; and Wilson, J.E. "Overview of Pay-Adjustment Factors for Asphalt Concrete Mixtures," *Transportation Research Record 821*, 1981, pp. 49-56.

In fall 1979, the Oregon State Highway Division and Oregon State University, with participation from the University of Washington, initiated a research project to study the impact of variations in material properties on asphalt pavement life. The study was aimed at developing a rational approach to assessing the effects of variations from specification limits so that a firm basis could be established for the development of pay factors. Analysis of the results indicates the following: (1) most State agencies will accept one or more property characteristics that are outside of specification tolerances, (2) most State agencies apply a pay factor for accepted material outside of specification tolerances, (3) only 26 percent of the State agencies consider their pay factor to be proportional to reduced pavement serviceability, (4) approximately half of the agencies consider pay-factor plans to be effective in encouraging compliance with specifications, and (5) there is a wide disparity in the pay-adjustment factors used by the different agencies.

Nachlas, J.A. and Kim, S-I. "Generalized Attribute Acceptance Sampling Plans," *Journal of Quality Technology*, Vol. 21, No. 1, 1989, pp. 32-40.

The imposition of acceptance criteria upon realizations of generalized sampling plans is described and has been shown to yield a large variety of new attribute acceptance sampling plans. The resulting generalized attribute acceptance sampling plans can be constructed to yield the same OC curve as that for conventional single sampling plans with lower expected total cost. An example is presented.

Nelson, L.S. "Factors for Confidence Limits on Standard Deviation," *Journal of Quality Technology*, Vol. 29, No. 4, 1997, pp. 485-487.

Tables are presented for easily determining both one-sided and two-sided confidence limits on standard deviations, assuming that the sample is random and from a normal population.

Patel, A.; Thompson, M.; Harm, E.; and Sheftick, W. "Developing QC/QA Specifications for Hot-Mix Asphalt Concrete in Illinois," *Transportation Research Record 1575*, 1997, pp. 66-74.

The Illinois DOT has recently undertaken a quality management program to improve the quality of construction, allow more innovation, and reduce the department's management of industry construction programs. The AC QC/QA program is a significant part of this quality management program. The Illinois DOT credits the success of the AC QC/QA program to gradual implementation, and contractor and industry involvement. In 1991, four projects were constructed under a newly developed QC/QA specification. After reviewing and evaluating feedback, the specification was revised for 1992. In 1992, 30 projects were constructed using the QC/QA specification and, in 1993, 65 projects used the specification. In 1994 and 1995, most projects over 225 Mg used the specification. In conjunction with this effort, an aggregate certification program was implemented. Training programs for contractor QC and aggregate certification were also implemented. Analysis of the data indicates an increase in the uniformity of the HMA, potentially leading to a 15-percent increase in fatigue life. The Illinois DOT is now examining the implications of developing ERS and PRS for the AC QC/QA program. In summer 1996, one QC/QA project was shadowed and evaluated based on newly developed ERS/PRS.

Pham, T.G. and Turkkan, N. "Bayes Binomial Sampling by Attributes With a General-Beta Prior Distribution," *IEEE Transactions on Reliability*, Vol. 41, No. 2, 1992, pp. 310-316.

In binomial sampling, the standard beta is frequently used as a prior because of its conjugate property for a sample from a Bernoulli distribution. The case where the prior is a general beta is examined. Practical advantages in the elicitation of expert opinion (to obtain a prior) and a convenient expression of the posterior are presented. A computer program on an IBM® personal computer (or a compatible computer) permits simple use of the general beta and is available from the authors. This program is used to solve the numerical examples, showing the advantages of using the general beta as a prior.

Schmitt, R.L.; Russell, J.S.; Hanna, A.S.; Bahia, H.U.; and Jung, G.A. "Summary of Current Quality Control/Quality Assurance Practices for Hot-Mix Asphalt Construction," *Transportation Research Record 1632*, 1998, pp. 22-31.

State highway agencies and contractors have been implementing QC/QA specifications in recent years to advance the quality of HMA construction. During their continued development, the attributes of these QC/QA specifications have varied among the States. This paper provides a compilation of state-of-the-art practices in

QC/QA for HMA construction and provides recommendations for State highway agencies and contractors when modifying or developing a QC/QA specification.

Seeds, S.B.; Basavaraju, R.; Epps, J.A.; and Weed, R.M. "Development of Performance-Related Specifications for Hot-Mix Asphalt Pavements Through WesTrack," *Transportation Research Record 1575*, 1997, pp. 85-91.

The primary purpose of the FHWA-sponsored WesTrack project is to further the development of PRS for HMA construction. This objective is being achieved, in part, through the accelerated loading of a full-scale test track facility in northern Nevada. Twenty-six HMA test sections that were constructed to meet the criteria set forth in a statistically based experimental design are providing performance data that will be used to improve (or develop new) pavement performance prediction relationships that better account for the effects that off-target values of asphalt content, air voids, and aggregate gradation have on such distress factors as fatigue cracking, permanent deformation, roughness, raveling, and tire-pavement friction. The concept of the planned PRS and how it will incorporate the modified pavement performance prediction models are described. The current plan for assessing contractor pay adjustments based on data collected from the as-constructed pavement is also discussed.

Semenov, V.A. "Quality Control in Highway Construction and Maintenance When the Measurement Parameters Are Highly Nonuniform," *Transportation Research Record 1126*, 1987, pp. 28-36.

An original method is presented for QC based on Weibull's law for the distribution of random quantities with variable parameters. The nomographs obtained for determining the extreme values of the parameters and the defectiveness index can be used for various cases in the statistical reduction of research results. The proposed method can be used for QC for both highly uniform data (normal distribution) and highly nonuniform data. The method described is widely used in the former Soviet Union for QC in the construction and maintenance of roads.

Sheng, Z. and Fan, D-Y. "Bayes Attribute Acceptance Sampling Plan," *IEEE Transactions on Reliability*, Vol. 41, No. 2, 1992, pp. 307-309.

An approach is reviewed for choosing a prior distribution for a Bayes attribute acceptance sampling plan. A prior distribution is chosen from confidence levels corresponding to classical lower confidence bounds. Where a Bayes plan is acceptable, the sample size can be reduced.

Soundararajan, V. "Maximum Allowable Percent Defective (MAPD) Single Sampling Inspection by Attributes Plan," *Journal of Quality Technology*, Vol. 7, No. 4, 1975, pp. 173-182.

Single sampling attribute plans indexed by maximum allowable percent defective (MAPD) are given. A table allowing the transitioning from one set of parameters to match the OC curve to other similar sets is given.

Soundararajan, V. and Arumainayagam, S.D. "Lot Sensitive Double Sampling Plan," *Communications in Statistics: Theory and Method*, Vol. 21, No. 10, 1992, pp. 2931-2948.

A sampling plan is derived for compliance testing, which provides consumer protection.

Suresh, K.K. and Ramkumar, T.B. "Selection of a Sampling Plan Indexed With Maximum Allowable Average Outgoing Quality," *Journal of Applied Statistics*, Vol. 23, No. 6, 1996, pp. 645-654.

A new concept of maximum allowable average outgoing quality (MAAOQ), which is the average outgoing quality at the inflection point, is introduced. The procedure for designing a single sampling plan indexed through the MAPD and MAAOQ is stated. Tables are constructed for the selection of parameters for the plan and parametric conversions are studied.

Suresh, R.P. and Ramamathan, T.V. “Acceptance Sampling Plans by Variables for a Class of Symmetric Distributions,” *Communications in Statistics: Simulation and Computation*, Vol. 26, No. 4, 1997, pp. 1379-1391.

The estimation of PD using a normal sampling plan is not appropriate when the assumption of normality is violated. A sampling plan based on a more general symmetrical family of distributions with the parameters estimated using the modified maximum likelihood (MML) procedures is presented. Some numerical studies are also presented, showing that the sampling plan works well for most of the symmetrical non-normal distributions.

Vlatas, D.A. and Smith, R.E. “Implications of Life-Cycle Performance Specifications,” *Transportation Research Record 1215*, 1989, pp. 25-30.

The professional, managerial, and legal implications of using life-cycle performance specifications are presented. Changes in the roles of the parties using life-cycle performance specifications are discussed. This approach can improve quality, reduce costs, and expedite the construction process. The basis of the process is the development of models of expected performance. These models will be used to predict whether the pavement will perform as required over the life of the project. Tests are performed at the end of construction to determine whether the expected performance is likely to be achieved. Adjustments in payment can be made, based on the performance model predictions. The ramifications of adopting life-cycle performance specifications are discussed.

Wasserman, G.S. “Matching Performance of Continuous Sampling Plans With Single Sampling Plans,” *Communications in Statistics: Simulation and Computation*, Vol. 19, No. 4, 1990, pp. 1303-1317.

Single sampling plans are widely used for appraising product quality. However, for situations where a continuous product flow exists, lot-by-lot demarcations may not exist, and it may be necessary to use alternate procedures, such as Continuous Sampling Plan 1 (CSP-1), for continuous processes. In this case, one would like to understand how the average performance of the continuous sampling procedures compares to the more commonly used single sampling plans. A model is devised that can be used to relate plan performance between single sample lot acceptance procedures and CSP-1. It is shown that it is generally not possible to match up performance based on OC expressions for the two plans. Instead, the plans are matched by equating expressions for π (p), the long run proportion of the product that is accepted, under both procedures. This is shown to be equivalent to matching properties on an average outgoing quality basis. Tables are generated that may be used to look up equivalent CSP-1 plans. The tables may also be used to match up plan performance around two target quality levels (p_o and p_1).

Weed, R.M. “Adjusted Pay Schedules: New Concepts and Provisions,” *Transportation Research Record 986*, 1984, pp. 32-38.

Shortly after the American Association of State Highway Officials (AASHO) Road Test had furnished a wealth of statistical data on pavement construction and performance, highway agencies began to use this data to develop ERS based on statistical concepts. These specifications usually included adjusted pay schedules, the development of which was sometimes quite arbitrary. More recently, attempts have been made to improve both the accuracy with which the pay schedules are established and the fairness with which they are administered. The rationale underlying several recent advances in the state-of-the-art is discussed. Included are the use of the principal liquidated damages to relate pay reductions to the anticipated monetary loss resulting from substandard work, the development of the crediting concept to overcome a basic inequity of many existing pay schedules, and the establishment of bonus provisions that provide additional incentive by awarding payment slightly in excess of the contract price for superior quality work.

Weed, R.M. “Analysis and Application of Correlated Compound Probabilities,” *Transportation Research Record 792*, 1981, pp. 49-53.

Many statistical applications require the calculation of compound probabilities and, frequently, the individual probabilities are not independent. The failure to recognize that a correlation exists in cases such as these has resulted in numerous errors in published literature. Although an exact analytical solution is not known, problems of this type can often be handled effectively by calculating lower and upper bounds for the desired probabilities. Bounds for both positively and negatively correlated cases are derived and then applied in the analysis of statistical acceptance procedures. The results of several computer simulation tests are presented to demonstrate the validity of the theoretically derived results.

Weed, R.M. "Composite Pay Equations: General Approach," *Transportation Research Record 1465*, 1994, pp. 9-15.

Highway construction specifications involving the acceptance testing of several different quality characteristics are sometime confusing and difficult to administer. A procedure is developed by which multiple quality measures may be combined in a rational manner in a single, composite pay equation. This approach is scientifically sound and may be applied to almost any construction specification for which a relationship between quality and performance is known or can be approximated. An example based on PCC pavement is presented to illustrate the practicality of this method.

Weed, R.M. "Computer-Assisted Random Sampling," *Transportation Research Record 1034*, 1985, pp. 140-152.

Many State transportation agencies use SQA specifications to govern construction work. A vital step in the application of these and other types of specifications is the selection of random samples to obtain a valid estimate of the quality received. Random sampling procedures are often tedious and time-consuming, but can be considerably simplified with computer assistance, either by using special forms generated by computer or by working directly at an interactive terminal. Examples of several applications are presented.

Weed, R.M. "Development of Air Voids Specification for Bituminous Concrete," *Transportation Research Record 1491*, 1995, pp. 33-39.

The New Jersey DOT has been using SQA specifications for various construction products since the 1960s. Throughout this period, there has been a continuing process leading to a better understanding of the operation and implementation of SQA procedures. The New Jersey DOT specification for air voids in bituminous concrete was one of the first to be developed and, as such, is a prime candidate for upgrading. A major change is to base the acceptance procedure on the PD rather than on the average of the test values in order to control both the level and the variability of the air voids in a statistically efficient way. Doing this required new definitions of AQL and rejectable quality level (RQL), and a reexamination of the adjusted pay schedule to be applied when other than AQL work is received. It was decided to use a bonus provision for superior quality, an approach that has worked well with other recently developed New Jersey DOT specifications. Another change is to use a continuous (equation-type) pay schedule to provide a smooth progression of payment as quality varies, thus avoiding potential disputes over measurement precision when a quality estimate falls just onto one side or the other of a boundary in a stepped pay schedule. The various developmental steps are described, including the construction of the OC curve to verify the performance of the specification and the field trials leading to its successful implementation.

Weed, R.M. "Development of Multicharacteristic Acceptance Procedures for Rigid Pavement," *Transportation Research Record 885*, 1982, pp. 25-36.

The manner in which the AASHTO design can be used to develop multicharacteristic acceptance procedures for rigid pavement is outlined. The AASHTO equation is used to compute both the expected load-bearing capacity based on the as-built characteristics of the pavement and the desired load-bearing capacity based on the design parameters. The ratio of these two values is then used to determine the appropriate pay adjustment, which may be either positive or negative. Sensitivity tests are performed to verify the reliability of this approach and computer simulation is used to demonstrate the effectiveness of several different acceptance procedures of this type. A secondary study is conducted to determine how the procedure based on the AASHTO equation compares with several other methods of treating multiple pay factors to obtain a single overall pay factor. Under the assumption that the AASHTO method is the fundamentally correct approach, the method of multiplying individual pay factors together is shown to be among the best of the other methods that were tested.

Weed, R.M. "Method to Establish Pay Schedules for Rigid Pavement," *Transportation Research Record 885*, 1982, pp. 18-24.

An equation is derived to compute the appropriate pay factor for any quality level of rigid pavement. The measure of quality used in this development is the estimated load-bearing capacity of the pavement, although the results may be applied to specifications based on other quality measures. The appropriate pay adjustment is considered to be the present worth of any expense or savings expected to occur in the future as the result of a departure from the specified level of quality and may be positive or negative. Sensitivity tests demonstrate that the method is reliable provided that the input variables are determined with reasonable accuracy. By using input values typical of a relatively urbanized area, this procedure indicates that a minimum pay factor of 60 percent is appropriate for the poorest quality work and a maximum pay factor of 115 percent is justified for work of truly superior quality. Additional factors are cited that, though unquantified, would tend to lower the minimum pay factor and raise the maximum pay factor. Finally, pay schedules are developed, the OC curves of which closely approximate the theoretically derived relationship.

Weed, R.M. "OCPLOT: Program to Generate Operating Characteristic Curves for Statistical Construction Specifications," *Transportation Research Record 1491*, 1995, pp. 18-26.

The performance of the Nation's highway system is inexorably linked to the quality of the design and the quality of the construction. To control the quality of the construction, transportation agencies have developed elaborate QA programs, most of which employ ERS that rely on statistical sampling and acceptance procedures to ensure that the work is done in accordance with the plans and specifications. Whether the acceptance procedure leads to a simple pass/fail decision or an adjustment in contract price, the proper design of such plans is critical to their performance. Poorly conceived plans may be totally ineffective or impractically severe, and both extremes have been found in published or proposed national standards. To encourage the design of plans that are both effective and fair, an interactive software program has been developed that enables the user to construct OC curves to analyze the performance of a wide range of acceptance plans. An example is presented to demonstrate the versatility of the program and the ease with which it can be applied.

Weed, R.M. "Practical Framework for Performance-Related Specifications," *Transportation Research Record 1654*, 1999, pp. 81-87.

As highway agencies moved away from the older prescription-type specifications and began to develop ERS and PRS, several different statistical measures of quality have been used. These include sample mean, PD or its complement, PWL, and AAD. The CI is yet another measure that has been proposed. What has not been undertaken during this developmental period is any sort of formal analysis to determine which, if any, of these measures accurately reflects the expected performance of the construction products to which they are applied. Specialized computer programs were developed to demonstrate the potential weaknesses of the current quality measures and to explore alternate approaches that may overcome these weaknesses. It was found that pay equations based on the mean and standard deviation computed from the sample can be tailored to closely match the value of the constructed product as estimated by life-cycle cost techniques. It is believed that this forms a practical starting point for the development of construction acceptance procedures more closely linked to quantified performance models.

Weed, R.M. "Quality Assurance Software for Personal Computer," *Transportation Research Record 1544*, 1996, pp. 116-124.

Demonstration Project 89 on Quality Management was created to provide guidance on the use of practical and effective procedures to ensure that the level of quality designed into the plans and specifications is accurately achieved in the constructed product. One part of this effort is the distribution of a software package consisting of several interactive programs developed for use on a personal computer. These extremely user-friendly programs make it possible to analyze both pass/fail and pay-adjustment acceptance procedures, construct OC curves, plot control charts, experiment with computer simulation, perform statistical comparisons of data sets, demonstrate the unreliability of decisions based on a single test result, and explore the effectiveness of stratified random sampling. This comprehensive software package provides highway engineers with the necessary tools for learning why some statistical procedures are inherently superior to others and how to incorporate this knowledge into fair and effective construction specifications.

Weed, R.M. "A Rational Method for Relating As-Built Quality to Pavement Performance and Value," *Transportation Research Record 1632*, 1998, pp. 31-39.

Highway construction specifications routinely use adjusted-payment provisions to award payment in proportion to the level of quality received. Work that is defined as acceptable is eligible for 100-percent payment, whereas work that fails to meet the desired quality level, but is not sufficiently deficient to warrant removal and replacement, typically receives some degree of pay reduction. There is not yet consistency in the practices regarding the magnitude of the pay adjustment that is judged as being appropriate for varying levels of as-built quality. There needs to be a method for relating as-built quality to expected performance, which can then be related to value by engineering economics procedures. The extension and refinement of earlier work are described and an example structured around one of the performance relationships in the AASHTO *Guide for Design of Pavement Structures* is provided.

Weed, R.M. "Revision of a Flawed Acceptance Standard," *Transportation Research Record 1056*, 1986, pp. 21-35.

A major revision of AASHTO Standard R9-84, Acceptance Sampling Plans for Highway Construction, has been completed. The primary goals were to correct a major conceptual error and to reduce the level of complexity. In this paper, the flaws in the original version are discussed, the basic changes that were made are described, and a significant addition to the new standard is presented. This addition is OC tables that enable the user to quickly and easily select acceptance plans that will provide the desired degree of QA. Computer simulation is used to demonstrate that single-limit variable OC curves are sufficiently accurate for most double-limit applications. Two examples are included to illustrate the use of the revised standard.

Weed, R.M. "Stratified Random Sampling From a Discrete Population," *Transportation Research Record 792*, 1981, pp. 41-44.

In the development of statistical acceptance procedures for products whose quantity is measured on a continuous scale by using units such as length, area, volume, or weight, QA engineers usually specify stratified random sampling plans to ensure a more uniform coverage of the product than is often achieved by pure random sampling. Stratified plans divide the total quantity of the product into an appropriate number of equal-sized sublots and require that a single random sample be taken from each. Not only is it desirable to develop an equivalent procedure for products that are measured in discrete units, but, in many cases, such a procedure will prove to be more convenient for continuous products that are delivered in discrete units, such as batches or truckloads. However, the development of such a procedure is not as straightforward as might be expected. The weaknesses of some of the more obvious approaches are discussed and then a method is presented that achieves the desired result.

Weed, R.M. and Strawderman, W.E. "Method to Exclude the Effect of Testing Error When Estimating the Percentage Defective of a Continuous Normal Population," *Transportation Research Record 792*, 1981, pp. 45-49.

The quality of a product is often characterized by the percentage of the population that falls outside of specific limits. Although established methods for estimating PD are accurate as far as the overall distribution of test results is concerned, part of the variability of this distribution is caused by the presence of testing errors that cause the PD of the product itself to be overestimated. A method has been developed to overcome this problem and computer simulation has been used to demonstrate that it is effective for situations in which the testing error is no larger than about one-half of the variability associated with the product. The results of several unsuccessful attempts to improve on this technique are also presented and described briefly.

Weingarten, H. "Confidence Intervals for the Percent Nonconforming Based on Variables Data," *Journal of Quality Technology*, Vol. 14, No. 4, 1982, pp. 207-210.

Sampling inspection by variables for percent nonconforming, such as with the use of MIL-STD-414, also provides the data for estimating the population percent nonconforming. A method is provided for obtaining a confidence interval for the population value by means of an OC curve. This OC curve is achieved from the data and is not that of the original sampling plan.

White, T.D. and Brown, E.R. “Statistical Quality Control Procedures for Airfield Pavement Materials and Construction,” *Transportation Research Record 652*, 1977, pp. 36-42.

The interaction of materials, construction, and the environment makes projections of pavement performance difficult. To increase the confidence with which these projections can be made, a simplified statistical QC plan for airfield pavement materials and construction was developed. Test results based on a specific number of samples can be extended with the desired confidence limits based on probability theory to evaluate an entire lot of material. This statistical QC plan evaluates pavement quality much better than do conventional specifications. It also describes how to handle materials or construction of borderline acceptability. When test results indicate that the material or construction meets the desired quality, it is accepted at 100-percent payment. If it is below the minimum requirements, it is rejected. When material or construction is found to be below the desired quality, but above the minimum requirement, it is accepted at a reduced price. The evaluation is done on a daily basis so that the engineer and the contractor know the level of acceptability as the project progresses.

Willenbrock, J.H. and Kopac, P.A. “Development of Price-Adjustment Systems for Statistically Based Highway Construction Specifications,” *Transportation Research Record 652*, 1977, pp. 52-58.

A methodology is presented that can be used to develop price-adjustment systems for use in statistically based highway construction specifications. Three approaches are proposed for the development of the price-adjustment system: (1) the serviceability approach, (2) the cost of production approach, and (3) the OC curve approach. The three approaches are discussed and compared, and their most appropriate applications are recommended. A fourth approach, the cost of QC approach, is also discussed, but is not fully developed because of the limited cost data available.

APPENDIX B: **SUMMARY OF AGENCY HMAC SPECIFICATIONS RECEIVED**

Alaska

Quality Control

The contractor is responsible for developing the mix design for the project, providing a QC plan that must be approved by the engineer prior to construction, and providing qualified personnel and equipment to conduct QC testing. The contractor must allow the engineer access to the laboratory to observe any and all testing procedures, calculations, test documentation, and plotting of the results. The frequency of the testing is determined by the contractor, but must be approved by the engineer prior to construction.

Mix design verification is evaluated for the following properties: Marshall stability, Marshall flow, percent voids in mix, dust-to-asphalt ratio, gradation, VFA, density, and VMA.

Quality Assurance

The engineer is responsible for all assurance testing. Assurance testing is conducted for the following characteristics: gradation, in-place density, and asphalt content. The State's assurance tests cannot be used for QC. The results of the tests will be made available to the contractor within 2 working days.

Acceptance

Acceptance is based on the results of the assurance tests. Acceptance is based on a lot size of 5000 tons (4535 Mg), with tests conducted on 500-ton (454 Mg) sublots. Acceptance of the material and any associated pay factors are determined using quality-level analysis.

Pay Factors

Pay factors are assigned for gradation, in-place density, and asphalt content. For density, the density pay factor (DPF) is determined by calculating the quality level of the lot and then selecting a pay factor from a table according to the sample size. The pay factor for gradation and asphalt content (composite pay factor (CPF)) is determined by the following equation:

$$CPF = [f_{\mathrm{19mm}}(PF_{\mathrm{19mm}}) + f_{\mathrm{12.5mm}}(PF_{\mathrm{12.5mm}}) + \ldots f_{ac}(PF_{ac})] \,/\, \sum f \qquad (37)$$

where: f factors are determined from the Weight Factors table below

After the DPF and CPF are known, the pay adjustment is determined by the following:

$$\text{Price Adjustment} = [(CPF \text{ or } DPF)^{*} - 1.00] \times (\text{Mg in lot}) \times (PAB) \qquad (38)$$

where: PAB = price-adjustment base (equal to \$33.00/Mg)

Weight Factors

Table 58. Alaska weight factors.

Gradation (mm)	*f* Factor
19.0	4
12.5	5
9.5	5
4.75	4
2.36	4
1.18	4
0.600	5
0.300	5
0.150	4
0.075	20
Asphalt Content	40

QC/QA Tests for Alaska

Table 59. Alaska QC/QA tests.

Test	Sampling Location	QC Testing Frequency by Contractor	Assurance Testing Frequency by Alaska DOT	Acceptance by Alaska DOT
Gradation (extracted)	Behind paver	Determined by contractor	One per sublot	Based on assurance testing
Asphalt content	Behind paver	Determined by contractor	One per sublot	Based on assurance testing
Density	Roadway	Determined by contractor	One per sublot	Based on assurance testing

Arkansas

Quality Control

The contractor is responsible for developing the mix design for the project, providing a QC plan that must be approved by the engineer prior to construction, and providing qualified personnel and equipment to conduct QC testing. The contractor must allow the engineer access to the laboratory to observe any and all testing procedures, calculations, test documentation, and plotting of results. The sampling location and testing frequency are determined by the contractor with the exception of gradation (one per 750 tons (680 Mg) minimum), and must be sufficient for control of the mix.

The contractor is responsible for QC for the following properties: gradation, asphalt content, air voids, VMA, Marshall stability, Maximum theoretical density, in-place density, and water sensitivity.

Mix design verification is evaluated for the following properties: gradation, Marshall stability, Marshall flow, air voids, VMA, density, asphalt content, dust-to-asphalt ratio, and water sensitivity.

Quality Assurance

The Arkansas State Highway and Transportation Department (AHDT) is responsible for all assurance sampling and testing. Testing is based on a lot size of 3000 tons (2720 Mg) with 750-ton (680-Mg) sublots. Assurance testing is conducted on gradation, asphalt content, air voids, VMA, Marshall stability, and in-place density.

Acceptance

The contractor's control test cannot be used for acceptance. Under the supervision of the engineer, the contractor must conduct acceptance testing on the materials. The sampling locations will be independent of the QC locations. Acceptance of the material is based on verification of the contractor's acceptance testing by the AHDT's assurance tests. Acceptance and adjustment is by lot. Lot size is 3000 tons (2720 Mg) with 750-ton (680 Mg) sublots. In addition, the engineer will check the smoothness of the pavement by using a rolling straightedge. Variations shall not exceed 5 millimeters (mm) for binder courses and 3 mm for surface courses. At least one pass shall be made for the full length of each lane. If the surface test is deficient, the surface must be corrected at the contractor's expense.

Pay Factors

Pay factors are assigned for in-place density, asphalt content, Marshall stability, air voids, and VMA. The pay factors for each property are determined according to the following schedule and the total reduction for the lot is the sum of the reductions for each property of the lot.

Asphalt Binder: The contract price of the entire lot will be reduced by 12 percent for each deviation outside the compliance limits, up to a maximum of three deviations. One deviation is 0.1 percentage point.

Stability: The contract price of the lot will be reduced by 10 percent for each deviation below the compliance limits, up to a maximum of five deviations. One deviation is 0.2 kilonewtons (kN) (45 pounds force (lbf)).

Air Voids: The contract price of the lot will be reduced by 10 percent for each deviation outside the compliance limits, up to a maximum of five deviations. One deviation is 0.1 percentage point.

VMA: The contract price of the lot will be reduced by 4 percent for each deviation outside the compliance limits, up to a maximum of 10 deviations. One deviation is 0.1 percentage point.

Density: The contract price of the lot will be reduced by 4 percent for each deviation outside the compliance limits, up to a maximum of 10 deviations. One deviation is 0.1 percentage points.

QC/QA Tests for Arkansas

Table 60. Arkansas QC/QA tests.

Test	Sampling Location	QC Testing Frequency by Contractor	Assurance Testing Frequency by AHDT	Acceptance Testing Frequency by Contractor
Gradation (extracted)	Behind paver	One per 750 tons	One per 750 tons	One per 750 tons
Asphalt content	Behind paver	Contractor's discretion	One per 3000 tons	One per sublot
Air voids	Behind paver	Contractor's discretion	One per 3000 tons	One per sublot
VMA	Behind paver	Contractor's discretion	One per 3000 tons	One per sublot
Maximum theoretical density	Laboratory sample	Contractor's discretion	NA	NA
In-place density	Behind paver	Contractor's discretion	One per 3000 tons	One per sublot
Water sensitivity	Laboratory sample	Contractor's discretion	NA	NA
Smoothness	Roadway	Contractor's discretion	NA	By engineer (one pass for full length of each lane)
Marshall stability	Behind paver	Contractor's discretion	One per 3000 tons	One per sublot

NA = Not applicable
1 ton = 0.907 Mg

Colorado

Quality Control
The contractor is responsible for developing the mix design for the project, providing a QC plan that must be approved by the engineer prior to construction, and providing qualified personnel and equipment to conduct QC testing. The contractor must allow the engineer access to the laboratory to observe any and all testing procedures, calculations, test documentation, and plotting of results.

All QC tests must be independent of assurance and acceptance tests. Acceptable testing standards in order of preference include Colorado procedures, AASHTO, and ASTM.

Quality Assurance
Colorado DOT ensures QA by conducting check tests to verify that the personnel and equipment conducting the QC tests and acceptance tests are within acceptable tolerances. If the check test is outside the allowable differences, production will stop and the problem will be identified and corrected.

Acceptance
Acceptance testing is the responsibility of the Colorado DOT. A minimum of 5 asphalt content, 3 gradation, and 10 density tests are required for each project. The contractor, under the direction of the engineer, obtains acceptance test samples. Acceptance is based on the State's tests only.

Pay Factors
Pay factors are assigned for gradation, density, and asphalt content. To determine the pay factors, quality-level analysis is used to obtain a quality level for a process with three or more test results. When the process has been completed, the number of random samples (*Pn*) that it contains will determine the computation of the element pay factor, based on the table and equation below. When the *Pn* is 3 to 9, or greater than 200, the element pay factor is computed using the equations designated in the table below. When the *Pn* is equal to or greater than 10, and less than 201, the element pay factor is computed using the following equation:

Pay Factor $(PF) = [(PF_1 + PF_2) / 2] + [(PF_2 + PF_3) / 2 - (PF_1 + PF_2) / 2] \times [(Pn_2 - Pn_x) / (Pn_2 - Pn_3)]$ (39)

where (when referring to the table below):

PF_1 = PF determined at the next lowest *Pn* equation using the process quality level (QL)
PF_2 = PF determined using the *Pn* equation shown for the process QL
PF_3 = PF determined at the next highest *Pn* equation using the process QL
Pn_2 = lowest *Pn* in the spread of values listed for the process *Pn* equation
Pn_3 = lowest *Pn* in the spread of values listed for the next highest *Pn* equation
Pn_x = actual number of test values in the process

Table 61. Colorado pay factor equations.

Pn	When the *Pn* is 3 to 9, or greater than 200, use the designated equation below to calculate the pay factor. When the *Pn* is equal to or greater than 10, and less than 201, use formula 1.	Maximum Pay Factor
3	$PF = 0.31177 + 1.57878(QL/100) - 0.84862(QL/100)^2$	1.025
4	$PF = 0.27890 + 1.15471(QL/100) - 0.73553(QL/100)^2$	1.030
5	$PF = 0.25529 + 1.48268(QL/100) - 0.67759(QL/100)^2$	1.030
6	$PF = 0.19468 + 1.56729(QL/100) - 0.70239(QL/100)^2$	1.035
7	$PF = 0.16709 + 1.58245(QL/100) - 0.68705(QL/100)^2$	1.035
8	$PF = 0.16394 + 1.55070(QL/100) - 0.65270(QL/100)^2$	1.040
9	$PF = 0.11412 + 1.63532(QL/100) - 0.68786(QL/100)^2$	1.040
10-11	$PF = 0.15344 + 1.50104(QL/100) - 0.58896(QL/100)^2$	1.045
12-14	$PF = 0.07278 + 1.62485(QL/100) - 0.65033(QL/100)^2$	1.045
15-18	$PF = 0.07826 + 1.55649(QL/100) - 0.56616(QL/100)^2$	1.050
19-25	$PF = 0.09907 + 1.43088(QL/100) - 0.45550(QL/100)^2$	1.050
26-37	$PF = 0.07373 + 1.41851(QL/100) - 0.41777(QL/100)^2$	1.055
38-69	$PF = 0.10586 + 1.26473(QL/100) - 0.29660(QL/100)^2$	1.055
70-200	$PF = 0.21611 + 0.86111(QL/100)$	1.060
≥ 201	$PF = 0.15221 + 0.92171(QL/100)$	1.060

For pay estimates, each individual element will have the average pay factor (PF_a), weighted by the quantities as follows:

$$PF_a = [M_1(PF_1) + M_2(PF_2) + \dots M_j(PF_j)] / \sum M \qquad (40)$$

where:
- Mj = quantity represented by the process
- PFj = process pay factor
- $\sum M$ = total quantity

QC/QA Tests for Colorado

Table 62. Colorado QC/QA tests.

Test	Sample Location	QC Testing Frequency by Contractor	Acceptance Testing Frequency by Colorado DOT	Check Testing Frequency by Colorado DOT
Asphalt content	Behind paver	One per 500 metric tons	One per 1000 metric tons	One per 10,000 metric tons
Dry gradation: #4, #8, #200	Cold feed	One per day	One per 2000 metric tons	One per 20,000 metric tons
In-place density	Roadway	One per 500 metric tons	One per 500 metric tons	One per 5000 metric tons

Idaho

Quality Control

The contractor is responsible for developing the mix design for the project, providing a QC plan that must be approved by the engineer prior to construction, and providing qualified personnel and equipment to conduct QC testing. The contractor must allow the engineer access to the laboratory to observe any and all testing procedures, calculations, test documentation, and plotting of results.

The contractor is responsible for QC for the following properties: gradation (3/4 inch (19 mm), 3/8 inch (9.5 mm), #8, and #200), sand equivalent, fracture count, asphalt content, and density. The lot size is 1 day's production, represented by at least four tests. If the day's production is represented by fewer than four tests, the tests will be combined with the following day's production to form one lot. One test must be conducted for every 750 tons (680 Mg) of production (never less than once per day). Asphalt content is the exception to this rule, with testing conducted as needed to control production.

Quality Assurance

Idaho DOT is responsible for all assurance testing. Assurance testing is conducted on gradation, sand equivalent, asphalt content, density, and fracture count. The QA testing frequency is two per lot, not to exceed two per day.

Acceptance

Acceptance testing is the responsibility of the contractor for the properties of gradation, sand equivalent, and fracture count. These acceptance tests are not to be used as part of the QC testing. The State is responsible for the acceptance testing of asphalt content and density. The lot size is 1 day's production, represented by at least four tests. If the day's production is represented by fewer than four tests, the tests will be combined with the following day's production to form one lot. One test must be conducted for every 750 tons (680 Mg) of production (never less than one per day).

Pay Factors

Pay factors are assigned for gradation, in-place density, and asphalt content. Quality-level analysis is used to determine the PWL for the lot. If any one pay factor for asphalt content, density, or gradation falls below 0.85, the pay factor for all quality characteristics represented by the same quantity will be determined by the lowest pay factor.

Once the individual gradation pay factors for the lots are determined, they are averaged on a weighted basis, considering the amount of material represented by each lot. This average gradation pay factor is then multiplied by 0.3 to obtain the composite pay factor for plant-mix aggregate (CPF_{PM}).

The individual pay factors for asphalt content are also averaged on a weighted basis to obtain the average asphalt content pay factor. This factor is then multiplied by 0.3 to obtain the composite pay factor for asphalt content (CPF_{AC}).

The same procedure is then performed to obtain the composite pay factor for density (CPF_{Dens}). The bonus to be paid or deducted is determined using the following equation:

$$B = (A)[(CPF_{PM} + CPF_{AC} + CPF_{Dens}) - 1](Q) \tag{41}$$

where: B = total bonus
A = unit bid price
Q = total quantity of plant mix accepted

QC/QA Tests for Idaho

Table 63. Idaho QC/QA tests.

Test	Sampling Location	QC Testing Frequency by Contractor	Assurance or Verification Testing Frequency by State	Acceptance
Gradation (washed) (¾ inch, ⅜ inch, #8, #200)	Crusher for QC, cold feed for acceptance	One per 750 tons, but not less than one per day	Two per lot, not to exceed two per day	By contractor: one per 750 tons, but not less than one per day
Asphalt content	Hot plant or roadway	As needed to control operation	Two per lot, not to exceed two per day	By State: one per 750 tons, but not less than one per day
Sand equivalent	Crusher for QC, cold feed for acceptance	One per 750 tons, but not less than one per day	Two per lot, not to exceed two per day	By contractor: one per 750 tons, but not less than one per day
Fracture count	Crusher for QC, cold feed for acceptance	One per 750 tons, but not less than one per day	Two per lot, not to exceed two per day	By contractor: one per 750 tons, but not less than one per day
Density	Roadway	One per 750 tons, but not less than one per day	Two per lot, not to exceed two per day	By State: one per 750 tons, but not less than one per day

1 inch = 25.4 mm, 1 ton = 0.907 Mg

Illinois

Quality Control

The contractor is responsible for developing the mix design for the project, providing a QC plan that must be approved by the engineer prior to construction, and providing qualified personnel and equipment to conduct QC testing. The contractor must allow the engineer access to the laboratory to observe any and all testing procedures, calculations, test documentation, and plotting of results.

The contractor is responsible for QC for the following properties: gradation (1/2 inch (13 mm), #4, #8, #30, and #200), BSG, MSG, air voids, and density. The contractor is responsible for providing split samples to the engineer for assurance testing.

Quality Assurance

The engineer is responsible for all assurance testing and acceptance of the material. The assurance tests are conducted on split samples taken by the contractor for QC testing. Additionally, the engineer will witness the sampling and splitting of these samples a minimum of twice a month and will retain these samples for assurance testing.

Acceptance

The contractor's control tests are used for acceptance if they are verified by the engineer's assurance tests. The testing frequencies vary, depending on the property being measured. Acceptance is based on the validation of the QC program, reviewing the contractor's control charts, and by the assurance tests for voids and density. Assurance testing is conducted on gradation (1/2 inch (13 mm), #4, #8, #30, and #200), BSG, MSG, air voids, and density.

Pay Factors

Pay factors are assigned for thickness and smoothness. Smoothness is determined using a California profilograph. The pay factor for smoothness is presented in the price-adjustment table for smoothness. The lot size for thickness testing shall be 1500 meters (m) (single lane). Each lot is divided into 10 sublots. One core will be taken from each sublot. Any pavement found to be less than 10-percent deficient may remain in place with no additional action required. After analyzing the cores, the PWL for the lot is determined. The following equations allow for the calculation of a bonus or penalty for thickness.

Thickness:

$$\text{Pay Factor Percentage} = 55 + 0.5PWL \tag{42}$$

$$\text{Total Payment} = PF\,[CUP(SQMPAVT - DEFPAVT)] \tag{43}$$

where:

PF = total pay factor
CUP = contract unit price
$SQMPAVT$ = square meters of pavement placed
$DEFPAVT$ = square meters of deficient pavement

Smoothness Pay Factors

Table 64. Illinois smoothness pay factors.

Profile Index for Entire Project (mm/km)	Percent of Unit Bid Price
≤ 3	103
> 3-6	102
> 6-8	101
Profile Index for 160-m Section (mm/km)	**Percent of Unit Bid Price**
> 8-160	100
> 160-175	98
> 175-190	96
> 190-205	94
> 205-220	92
> 220-235	90
> 235	Corrective work required

QC/QA Tests for Illinois

Table 65. Illinois QC/QA tests.

Test	Sampling Location	QC Testing Frequency by Contractor	Assurance Testing Frequency by Illinois DOT	Acceptance by Illinois DOT
Dry gradation (½ inch, #4, #8, #30, #200)	Cold Feed	One per ½ day	≥ 10% of QC	Based on validation of QC tests by assurance tests
MSG	Plant (hot mix)	One per ½ day for first 2 days, then one per day	≥ 20% of QC	Based on validation of QC tests by assurance tests
BSG	Plant (hot mix)	One per ½ day for first 2 days, then one per day	≥ 20% of QC	Based on validation of QC tests by assurance tests
Air voids	Plant (hot mix)	One per ½ day for first 2 days, then one per day	≥ 20% of QC	Based on validation of QC tests by assurance tests
Thickness	Roadway	One per sublot	NA	Based on contractor's tests
Asphalt content	Plant (hot mix)	One per ½ day	≥ 20% of QC	Based on validation of QC tests by assurance tests
Smoothness	Roadway	One pass per lane per 1 day's paving (1000 ft)	NA	Based on contractor's tests
Density	Roadway	One per ½ mi	≥ 20% of QC	Based on validation of QC tests by assurance tests

1 inch = 25.4 mm, 1 foot (ft) = 0.305 meter (m), 1 mile (mi) = 1.61 kilometers (km)

Iowa

Quality Control

The contractor is responsible for developing the mix design for the project, providing a QC plan that must be approved by the engineer prior to construction, and providing qualified personnel and equipment to conduct QC testing. The contractor must allow the engineer access to the laboratory to observe any and all testing procedures, calculations, test documentation, and plotting of results.

The contractor is responsible for QC for the following properties: laboratory density, laboratory voids, asphalt content, gradation, MSG, and field density. The contractor must provide split samples to the engineer for assurance testing. One of each control test must be conducted on the first 500 tons (450 Mg) and the remaining day's production is divided into three equal sublots and tested once per sublot.

Quality Assurance

Iowa DOT is responsible for all assurance testing and acceptance of the material. Assurance testing frequency is at the discretion of the engineer, but must be at least 10 percent of the contractor's QC testing, except for density and thickness, which require a minimum of seven cores per lot. Properties tested include gradation, laboratory density, laboratory voids, asphalt content, MSG, air voids, thickness, and in-place density.

Acceptance

The contractor's control test will be used for acceptance when the test results are verified by Iowa DOT's assurance tests.

Pay Factors

Pay factors are assigned for in-place density and thickness, based on a quality index for these two properties. The pay factors are based on square yards per lot. The pay factors for density are determined using the following tables:

Pay Factors for Density

Table 66. Iowa pay factors for density.

Quality Index	Percentage of Full Payment
0.73	100
0.40-0.72	95
0.00-0.39	85
All negative values	75 maximum or remove

Pay Factors for Thickness

Table 67. Iowa pay factors for thickness.

Quality Index	Percentage of Payment Previously Adjusted for Density
0.35+	100
0.14-0.34	95
0.00-0.13	85
All negative values	75 maximum or remove

QC/QA Tests for Iowa

Table 68. Iowa QC/QA tests.

Test	Sampling Location	QC Testing Frequency by Contractor	Assurance Testing Frequency by Iowa DOT	Acceptance by Iowa DOT
Gradation (extracted)	Behind paver	One per day	As needed for QC verification	Based on QC tests verified by assurance tests
Laboratory density	Plant mix (laboratory)	One per sublot	As needed for QC verification	Based on QC tests verified by assurance tests
Laboratory voids	Plant mix (laboratory)	One per sublot	As needed for QC verification	Based on QC tests verified by assurance tests
Asphalt content	Behind paver	One per sublot	As needed for QC verification	Based on QC tests verified by assurance tests
MSG	Plant mix (laboratory)	One per sublot	As needed for QC verification	Based on QC tests verified by assurance tests
Thickness	Roadway	One per sublot	Seven cores per lot	Based on QC tests verified by assurance tests
Air voids	Roadway	One per sublot	As needed for QC verification	Based on QC tests verified by assurance tests
In-place density	Roadway	One per sublot	Seven cores per lot	Based on QC tests verified by assurance tests

Maryland

Quality Control
The contractor is responsible for developing the mix design for the project, providing a QC plan that must be approved by the engineer 30 days prior to construction, and providing qualified personnel and equipment to conduct QC testing. The contractor must allow the engineer access to the laboratory to observe any and all testing procedures, calculations, test documentation, and plotting of results.

The contractor is responsible for QC for the following properties: gradation, density, asphalt content, MSG, VTM, VMA, Marshall flow smoothness, and Marshall stability. The contractor must provide split samples to the engineer for assurance testing.

Quality Assurance
Maryland DOT is responsible for all assurance testing. Assurance testing is conducted on split samples provided by the contractor. Properties measured include gradation, density, asphalt content, MSG, VTM, VMA, Marshall flow smoothness, and Marshall stability. Additional assurance samples may be taken and tested if deemed necessary by the engineer. The lot size for all properties is 3000 tons (2720 Mg) with 1000-ton (910-Mg) sublots, except for density, which is 1000 tons (910 Mg) with 200-ton (180-Mg) sublots.

Acceptance
Independent tests on split samples, observing contractor's testing, monitoring control charts, additional sampling as deemed necessary by the engineer, and monitoring the contractor's adherence to the QC plan provides acceptance. The lot size for all properties is 3000 tons (2720 Mg) with 1000-ton (910-Mg) sublots, except for density, which is 1000 tons (910 Mg) with 200-ton (180 Mg) sublots.

Pay Factors
Pay factors are assigned for gradation, in-place density, asphalt content, and smoothness. A pay factor for density and a composite pay factor for binder content and gradation is computed using quality-level analysis—the standard deviation method to determine the PWL and the associated pay factor. The lot payment is then determined using the following equation:

$$\text{Lot Payment} = (CUP)(PF_{Dens})(CPF_{GAC})(\text{tonnage}) \qquad (44)$$

where: CUP = contract unit price
PF_{Dens} = density pay factor
CPF_{GAC} = composite pay factor for gradation and density

Smoothness is measured by a computerized profilograph that takes measurements in 0.1-mi (0.16-km) increments and must conform to the following:

1. Single-lift construction with a wedge and leveling on tangent alignment and pavement on horizontal curves having a centerline radius of 2000 ft (610 m) or more shall have a profile index of ≤ 10 inches/mi (158 mm/km).
2. Multiple-lift construction, with or without a wedge and leveling on tangent alignment and pavement on horizontal curves having a centerline radius of 2000 ft (610 m) or more shall have a profile index of ≤7 inches/mi (110 mm/km).
3. Either of the above on horizontal curves having a centerline radius of curve of 1000 ft (305 m) or more, but less than 2000 ft (610 m), and pavement within the superelevation transition of such curves shall have a profile index of ≤ 12 inches/mi (190 mm/km).
4. Single-lift construction without wedge and leveling placed on a tangent alignment pavement after grinding shall have a profile index of ≤ 12 inches/mi (190 mm/km).
5. Single-lift construction without wedge and leveling placed on pavement shall have a profile index of ≤ 15 inches/mi (240 mm/km).

Areas that are accepted at a reduced price for a profile index deficiency will be adjusted by the factors shown in the following table in conformance with the procedures specified in items 1 through 5 above:

Table 69. Maryland profile index adjustment (normal projects).

Profile Index Exceeds Specification (inches/mi per 0.1-mi section)	Percentage of Payment Unit Bid Price
0.1-1.0	98
1.1-2.0	96
2.1-3.0	94
3.1-4.0	92
4.1-5.0	90
≥ 5.1	Corrective work required

1 inch/mi = 16 mm/km, 1 mi = 1.61 km

This adjustment is made at the end of paving and is based on the overall average profile index for the project.

Areas that are accepted on incentive projects shall use the following adjustment pay factor schedule and shall be in conformance with items 1 through 5 above.

Table 70. Maryland profile index adjustment (incentive projects).

Final Average Profile Index				Percentage of Pavement Unit Price
(5)	(3,4)	(1)	(2)	
0.0-1.9	0.0-1.9	0.0-1.9	0.0-1.9	105
2.0-2.9	2.0-2.9	2.0-2.9	2.0-2.9	104
3.0-3.9	3.0-3.9	3.0-3.9	3.0-3.9	103
4.0-15.0	4.0-10.0	4.0-10.0	4.0-7.0	100

QC/QA Tests for Maryland

Table 71. Maryland QC/QA tests.

Test	Sampling Location	QC Testing Frequency by Contractor	Assurance Testing Frequency by Maryland DOT	Acceptance by Maryland DOT
Density	Roadway	Five cores per 1000-ton lot	Three cores per 6000 tons	Based on verification of QC by QA
Gradation	Cold feed	One per 1000 tons or per day	One per lot	Based on verification of QC by QA
Asphalt content	Plant (hot mix) or behind paver	One per 1000 tons or per day	One per lot	Based on verification of QC by QA
MSG	Plant (hot mix) or behind paver	One per 1000 tons or per day	One per lot	NA
VTM	Plant (hot mix) or behind paver	One per 1000 tons or per day	One per lot	NA
VMA	Plant (hot mix) or behind paver	One per 1000 tons or per day	One per lot	NA
Flow	Plant (hot mix) or behind paver	One per 1000 tons or per day	One per lot	NA
Stability	Plant (hot mix) or behind paver	One per 1000 tons or per day	One per lot	NA
Smoothness	Roadway	0.1-mi segments	Randomly at engineer’s discretion	Based on verification of QC by QA

NA = Not applicable, 1 mi = 1.61 km, 1 ton = 0.907 Mg

Michigan

Quality Control

The contractor is responsible for developing the mix design for the project, providing a QC plan that must be approved by the engineer prior to construction, and providing qualified personnel and equipment to conduct QC testing. The contractor must allow the engineer access to the laboratory to observe any and all testing procedures, calculations, test documentation, and plotting of results.

The contractor is responsible for QC for the following properties: gradation (#4, #30, and #200), air voids, VMA, TMD, binder content, and crushed particle content.

Quality Assurance

Michigan DOT is responsible for all assurance testing and acceptance of the material. QA testing on loose mix and on compacted mix will be done at the field laboratory.

Acceptance

Acceptance is based on verification of the contractor's QC tests. If the engineer's tests and the contractor's tests are within allowable tolerances for the job mix formula, the material is considered to be acceptable. Testing is conducted on equal sublot sizes with a maximum size of 1200 tons (1090 Mg). Five consecutive sublots constitute a lot.

Pay Factors

Pay factors are assigned for gradation, TMD, binder content, crushed particle count, density, and VMA. The criteria for each of the pay factors are not available at this time. However, the negative price adjustments for the bituminous mixture contract unit price is cumulative according to the following equation:

$$\text{Price Adjustment} = (CUP)(\text{Pavement density} + \text{Mixture properties} + \text{Failure to suspend operations}) \quad (45)$$

where: CUP = contract unit price
Mixture properties = gradation, TMD, binder content, crushed particle count, and VMA

QC/QA Tests for Michigan

Table 72. Michigan QC/QA tests.

Test	Sampling Location	QC Testing Frequency by Contractor	Assurance Testing Frequency by Michigan DOT	Acceptance by Michigan DOT
TMD	Plant	One per sublot	One per lot	Based on acceptable tolerances with JMF
Gradation (dry): #4, #30, #200	Cold feed	One per day	One per lot	Based on acceptable tolerances with JMF
Crushed particle content	Cold feed	One per day	One per lot	Based on acceptable tolerances with JMF
Binder content	Plant	One per sublot	One per lot	Based on acceptable tolerances with JMF
Air voids	Plant	One per sublot	One per lot	Based on acceptable tolerances with JMF
VMA	Plant	One per sublot	One per lot	Based on acceptable tolerances with JMF
In-place density	Roadway	Three per sublot	One per lot	Based on acceptable tolerances with JMF

Minnesota

Quality Control

The contractor is responsible for developing the mix design for the project, providing a QC plan that must be approved by the engineer prior to construction, and providing qualified personnel and equipment to conduct QC testing. The contractor must allow the engineer access to the laboratory to observe any and all testing procedures, calculations, test documentation, and plotting of results.

The contractor is responsible for QC for the following properties: gradation, asphalt content, Marshall BSG, MSG, air voids, VMA, TSR, crushed count, and moisture content. Testing frequency is determined by taking the day's planned production, dividing by 1000, then rounding that figure to the next whole number.

Quality Assurance

Minnesota DOT is responsible for all assurance testing and acceptance of the material. At least one set of assurance tests is run on one set of production tests per day. Assurance is accomplished by testing, observing the contractor's QC sampling and testing, taking additional samples, and monitoring QC control charts.

Acceptance

The contractor's control test cannot be used for acceptance. Acceptance is based on Minnesota DOT's assurance tests. Pavement smoothness is the exception. The contractor's testing for smoothness is observed by the engineer and is used for acceptance and pay adjustments.

Pay Factors

Pay factors are assigned for gradation, VMA, asphalt content, density, profile index, and air voids. Pay factors for gradation, VMA, asphalt content, and air voids are determined from the following table, with the lowest single payment applied:

Payment Schedule

Table 73. Minnesota payment schedule.

Item	Percentage of Payment
Gradation	90
VMA	85
Asphalt content	85
Air voids (individual)	70

The pay factor for in-place density is determined on a lot-by-lot basis, with the number of lots determined by the following table:

Determination of Lots for Density

Table 74. Minnesota determination of lots for density.

Daily Production (tons)	Lots
0-300	1
301-600	2
601-1000	3
1001-1600	4
1601-3600	5
3601-5000	6
5000+	7

1 ton = 0.907 Mg

When the density of a lot of compacted mixture is less than the specified minimum, payment will be made at an adjusted price as specified in the following table:

Adjusted Payment Schedule for Maximum Density (Disincentive)

Table 75. Minnesota adjusted payment schedule for maximum density (disincentive).

Percentage of Density Below Specified Minimum	Payment Factor (percentage of contract price)
0.1-1.0 inclusive	98
1.1-1.5 inclusive	95
1.6-2.0 inclusive	91
2.1-2.5 inclusive	85
2.6-3.0 inclusive	70
> 3.0	Remove and replace

An incentive is available for density exceeding the specified minimum according to the following table:

Adjusted Payment Schedule for Maximum Density (Incentive)

Table 76. Minnesota adjusted payment schedule for maximum density (incentive).

Percentage of Density Above Specified Minimum	Pay Factor (percentage of contract price)
0.1-1.0 inclusive	100
1.1-1.5 inclusive	102
> 1.5	104

The profile index is determined using a California-type profilograph. Profiles will be made for segments of 0.1 km.

The following table determines the pay factor for smoothness:

Payment Schedule for Smoothness

Table 77. Minnesota payment schedule for smoothness.

mm per km per 0.1-km segment	Dollars per Segment
0-13	110
14-25	90
26-38	65
39-79	0
80-92	-90
93-105	-180
106-118	-360
> 118	Corrective action

QC/QA Tests for Minnesota

Table 78. Minnesota QC/QA tests.

Test	Sampling Location	QC Testing by Contractor	Assurance Testing Frequency by Minnesota DOT	Acceptance by Minnesota DOT
Gradation (extracted)	Truck or paver	One per 2200 tons	One per 2200 tons	Based on assurance test results
Asphalt content	Truck or paver	One per 1000 tons	One per day	Based on assurance test results
VMA	Truck or paver	Two per day	One per day	Based on assurance test results
TSR	Truck or paver	One per 11,000 tons	Engineer's discretion	Based on assurance test results
Density	Roadway	Three per lot	One per day	Based on assurance test results
Air voids	Truck or paver	Two per day	One per day	Based on assurance test results
Crushed count	Cold feed	Two per day for 2 days then one per day	One per day	Based on assurance test results
MSG	Truck or paver	Two per day	One per day	Based on assurance test results
Smoothness	Roadway	One pass per lane segment (0.1 km)	Observation of QC	Based on QC testing as observed by the engineer
BSG	Truck or paver	Two per day	One per day	Based on assurance test results

1 ton = 0.907 Mg

Montana

Quality Control

The contractor is responsible for developing the mix design for the project, providing a QC plan that must be approved by the engineer prior to construction, and providing qualified personnel and equipment to conduct QC testing. The contractor must allow the engineer access to the laboratory to observe any and all testing procedures, calculations, test documentation, and plotting of results.

The contractor is responsible for QC for the following properties: gradation, density, asphalt content, and fractured faces.

Quality Assurance

Montana DOT is responsible for all assurance sampling and testing. Assurance testing is conducted on the properties evaluated for QC. Assurance is conducted on a sublot size of 600 tons (540 Mg).

Acceptance

The contractor's control test cannot be used for acceptance. Acceptance is determined by evaluating the QA tests for compliance with the JMF tolerances for gradation, in-place density, and asphalt content. Quality incentives and disincentives are established by determining if the mix is within the allowable tolerances of the JMF. Acceptance is based on a 3000-ton (2720-Mg) lot.

Pay Factors

Pay factors are assigned for gradation, in-place density, and asphalt content. The factors are determined according to the following equations and the Price Reduction Factors table. All of the individual test results in the lot for the element to be evaluated will be averaged and the percentage of price reduction for the lot will be determined by the applicable equation:

1. The equation $P = (Xn + aR - Tu) \times F$ will be used if a maximum limit only is specified, or when the average of several test values is above the midpoint of a specified band or above a job mix target value. (46)

2. The equation $P = (TL + aR - Xn) \times F$ will be used if a minimum limit only is specified, or when the average of several test values is below the midpoint of a specified band or below a job mix target value. (47)

where:

P	=	percent reduction in the contract price
Xn	=	average of several test values from samples from the lot. (n = sample size)
a	=	variable factor used as n changes: $n = 3$, $a = 0.45$; $n = 4$, $a = 0.38$; $n = 5$, $a = 0.33$; $n = 6$, $a = 0.30$; and $n = 7$, $a = 0.28$
R	=	difference between the highest and lowest values in the group of values
Tu	=	upper tolerance limit
TL	=	lower tolerance limit
F	=	price reduction factor from the following table:

Price Reduction Factors

Table 79. Montana price reduction factors.

Element	*f* Factor
100% sieve	1
½-inch sieve and larger	1
#100 sieve to ⅜-inch sieve inclusive, except 100% sieve size	Cover material: 2 All other aggregate: 3
#200	Cover material: 3 All other aggregate: 6
Density	12
Fractured faces	2

1 inch = 25.4 mm

If $P < 3$ or a negative quantity, the lot is accepted as being in conformance. If one or more elements for a contract item show a positive P-value, the positive values are added and the resulting sum is used to determine whether the lot is in conformance. If $P = 3$ to 25, corrective action is required or the lot is accepted at a reduced price. If $P = 25+$, remove and replace. An incentive pay factor of 1.05 is allowed when the aggregate gradation for #4, #40, and #200 sieves is not more than one-half the allowable tolerance from the JMF. Additionally, a pay factor of 1.05 will be applied when the average density of the lot is from 97 percent to 98 percent, inclusive, from the target field Marshall density and the range is 3 or less.

QC/QA Tests for Montana

Table 80. Montana QC/QA tests.

Test	Sampling Location	QC Testing Frequency by Contractor	Assurance Testing Frequency by Montana DOT	Acceptance by Montana DOT
Gradation (dry)	Cold feed	One per 600 tons	One per 600 tons	Based on QA testing
Density	Roadway	One per 600 tons	One per 600 tons	Based on QA testing
Asphalt content	Behind paver	One per 600 tons	One per 600 tons	Based on QA testing
Fractured faces	Cold feed	One per 600 tons	One per 600 tons	NA

NA = Not applicable
1 ton = 0.907 Mg

Nebraska

Quality Control
The contractor is responsible for developing the mix design for the project, providing a QC plan that must be approved by the engineer prior to construction, and providing qualified personnel and equipment to conduct QC testing. The contractor must allow the engineer access to the laboratory to observe any and all testing procedures, calculations, test documentation, and plotting of results.

The contractor is responsible for QC for the following properties: gradation, BSG, ESG, binder content, TMD, air voids, VMA, and density.

Mix design verification is evaluated for the following properties: number of Marshall blows, Marshall stability, Marshall flow, percent voids in mix, dust-to-asphalt ratio, binder content, VMA, crushed face, TMD, BSG, and ESG.

Quality Assurance
Nebraska Department of Roads (DOR) is responsible for all assurance testing and acceptance of the material. All assurance test samples are split samples obtained from the contractor under the direction of the Nebraska DOR. From each sample, three Marshall samples are made and tested. Cores for density testing are obtained by Nebraska DOR. Assurance testing is conducted on the properties evaluated for QC.

Acceptance
Acceptance is based on verification of the contractor's QC tests for air voids, VMA, gradation, BSG, and in-place density by Nebraska DOR's assurance tests. If the results are verified, Nebraska DOR's tests are used for acceptance.

Pay Factors
A pay factor is assigned for density. The appropriate pay factor is then multiplied by the contract unit price. The pay factor for density is determined according to the following tables:

Density of Asphalt Concrete (First Lot)

Table 81. Nebraska density of asphalt concrete (first lot).

Average Density (Five Samples)	Pay Factor
> 90.0	1.00
> 89.5-90.0	0.95
> 89.0-89.5	0.70
≤ 89.0	0.40 or reject

Density of Asphalt Concrete (Subsequent Lots)

Table 82. Nebraska density of asphalt concrete (subsequent lots).

Average Density (Five Samples)	Pay Factor
> 92.4	1.00
> 91.9-92.4	0.95
> 91.4-91.9	0.90
> 90.9-91.4	0.85
> 90.4-90.9	0.80
> 89.9-90.4	0.70
≤ 89.9	0.40 or reject

QC/QA Tests for Nebraska

Table 83. Nebraska QC/QA tests.

Test	Sampling Location	QC Testing Frequency by Contractor	Assurance Testing Frequency by Nebraska DOR	Acceptance by Nebraska DOR
Marshall test	Plant mix	One per 1100 tons	One per 1 day's production	Based on verification of contractor's QC tests by Nebraska DOR's QA tests
BSG	Plant mix	One per 1100 tons	One per 1 day's production	Based on verification of contractor's QC tests by Nebraska DOR's QA tests
TMD	Plant mix	One per 1100 tons	One per 1 day's production	NA
ESG	Plant mix	One per 1100 tons	One per 1 day's production	NA
Asphalt content	Plant mix or behind paver	One per 1100 tons	One per 1 day's production	Based on verification of contractor's QC tests by Nebraska DOR's QA tests
Air voids	Plant mix or behind paver	One per 1100 tons	One per 1 day's production	Based on verification of contractor's QC tests by Nebraska DOR's QA tests
VMA	Plant mix	One per 1100 tons	One per 1 day's production	Based on verification of contractor's QC tests by Nebraska DOR's QA tests
Density	Roadway	Five per 2750 tons	Five per 1 day's production	Based on verification of contractor's QC tests by Nebraska DOR's QA tests
Gradation	Cold feed	One per 1100 tons	One per 1 day's production	Based on verification of contractor's QC tests by Nebraska DOR's QA tests

NA = Not applicable
1 ton = 0.907 Mg

Nevada

Quality Control

The contractor is responsible for developing the mix design for the project, providing a QC plan that must be approved by the engineer prior to construction, and providing qualified personnel and equipment to conduct QC testing. The contractor must allow the engineer access to the laboratory to observe any and all testing procedures, calculations, test documentation, and plotting of results.

The contractor must conduct QC tests on the following properties: gradation, liquid limit and plastic index, in-place density, absorption, fractured faces, theoretical MSG, moisture content, binder temperature, and mix temperature. The contractor is responsible for providing split samples to the engineer for verification and referee testing.

Nevada DOT is responsible for conducting the QC tests for Hveem stability, ITS, and TSR.

Quality Assurance

The contractor is responsible for all assurance testing. The testing frequencies vary, depending on the property being measured. Nevada DOT conducts verification testing to determine if the contractor's assurance test results are within acceptable limits. If the test values are not within the tolerances, referee testing is conducted.

Acceptance

Acceptance is based on the results of the QC tests, assurance tests, verification tests, and referee tests (if required). A production lot is four equal sublots. The sublot size is agreed upon by the contractor and the engineer prior to construction and may be 250, 500, or 750 tons (230, 450, or 680 Mg).

Pay Factors

Pay factors are assigned for gradation, in-place density, asphalt content, and profile index. The pay factors for each of these properties is determined from the following tables:

Pay Factors for Profile Index

Table 84. Nevada pay factors for profile index.

Pay Factor	mm/km		
	Type A	Type B	Type C
1.05	0-45	0-80	0-85
1.04	46-60	81-94	86-110
1.02	61-70	95-100	111-140
1.00	71-80	101-110	141-160
0.98	81-95	111-125	161-165
0.96	96-110	126-140	166-175
0.94	111-125	141-160	176-180
0.92	126-140	161-175	181-190
0.90	141-160	176-190	—
Corrective work	> 160	> 190	> 190

Pay Factors for Gradation

Table 85. Nevada pay factors for gradation.

Pay Factor	Mean Absolute Deviation From JMF Target	
	Pass 2 mm	Pass 75 μm
1.05	0.00-0.99	0.00-0.50
1.02	1.00-1.90	0.51-0.90
1.00	1.91-3.00	0.91-1.50
0.95	3.01-4.00	1.51-2.00
0.90	4.01-5.00	2.01-2.50
0.85	5.01-6.00	2.51-3.00
0.80	6.01-7.00	3.01-3.50
0.75	7.01-8.00	3.51-4.00
0.70	8.01-9.00	4.01-4.50
Remove	> 9.00	> 4.50

Pay Factors for Asphalt Content

Table 86. Nevada pay factors for asphalt content.

Pay Factor	Mean Absolute Deviation From JMF Target
1.05	0.00-0.19
1.02	0.20-0.24
1.00	0.25-0.30
0.95	0.31-0.35
0.90	0.36-0.40
0.85	0.41-0.45
0.80	0.46-0.50
0.75	0.51-0.60
0.70	0.61-0.65
Remove	> 0.65

Pay Factors for In-Place Air Voids

Table 87. Nevada pay factors for in-place air voids.

Pay Factor	Measured Air Voids (average of five locations per sublot)
1.05	4.0-5.9
1.02	6.0-6.9
1.00	7.0-8.0
0.95	3.5-3.9 or 8.1-9.5
0.90	3.0-3.4 or 9.6-10.0
0.80	2.6-2.9 or 10.1-11.0
0.70	11.1-12.0

Note: If the average in-place air void for a sublot is less than 2.5 or greater then 12.0, the entire sublot must be removed.

Once the pay factors have been determined, the pay calculation is as follows:

A = HMA set price × Placement quantity × Production pay factor
B = HMA set price × Placement quantity tested for air voids × Placement pay factor
C = HMA set price × Placement quantity tested for rideability × Ride pay factor

The total pay quantity (*TPQ*) shall be based on the applicable pay adjustment factors for production, placement, and ride quality:

For production only: $TPQ = A$
For production and placement: $TPQ = 0.40A + 0.60B$
For production, placement, and ride quality: $TPQ = 0.30A + 0.40B + 0.30C$

The incentive or disincentive pay amount is the difference between the standard pay quantity (SPQ) and the TPQ.

QC/QA Tests for Nevada

Table 88. Nevada QC/QA tests.

Test	Sample Location	QC Testing Frequency	Assurance Testing Frequency by Nevada DOT	Verification Testing Frequency by Nevada DOT
Gradation (except 2 mm and 75 µm)	Cold feed	One per sublot by contractor	NA	1 per 12 sublots by Nevada DOT
Gradation (2 mm and 75 µm)	Behind paver	NA	One per sublot by contractor	1 per 12 sublots by Nevada DOT
Asphalt content	Behind paver	NA	One per sublot by contractor	1 per 12 sublots by Nevada DOT
HVEEM stability	Behind paver	One per eight sublots by Nevada DOT	NA	NA
Theoretical MSG	Behind paver	Average of two random tests per lot by contractor	NA	1 per 12 sublots by Nevada DOT
Moisture content	Behind paver	One per four sublots by contractor	NA	1 per 12 sublots by Nevada DOT
Binder temperature	Plant	Continuous by contractor	NA	Continuous by Nevada DOT
Mix temperature	Truck bed	Continuous by contractor	NA	Continuous by Nevada DOT
In-place air voids	Roadway	NA	Five per sublot by contractor	1 per 12 sublots by Nevada DOT
Liquid limit	Stockpile	One per four sublots by contractor	NA	1 per 12 sublots by Nevada DOT
Plastic index	Stockpile	One per four sublots by contractor	NA	1 per 12 sublots by Nevada DOT
Fractured faces	Stockpile	One per four sublots by contractor	NA	One per four sublots by Nevada DOT
Absorption	Stockpile	One per four sublots by contractor	NA	1 per 12 sublots by Nevada DOT
ITS	Behind paver	1 per 20 sublots by Nevada DOT	NA	NA
Smoothness	Roadway	One pass per lane of travel per 0.1 km	One pass per lane of travel per 0.1 km	One pass per lane per lot
TSR	Behind paver	1 per 20 sublots by Nevada DOT	NA	NA

NA = Not applicable

North Dakota

Quality Control

The contractor must use a mix design developed by North Dakota DOT. It is the contractor's responsibility to meet the design requirements of the JMF. In addition, the contractor is responsible for providing a QC plan that must be approved by the engineer prior to construction and for providing qualified personnel and equipment to conduct QC testing. The contractor must allow the engineer access to the laboratory to observe any and all testing procedures, calculations, test documentation, and plotting of results.

The contractor is responsible for QC for the following properties: gradation, plastic index, lightweight aggregate, fractured faces, MSG, BSG, air voids, asphalt content, and density. The contractor must provide split samples for the engineer to conduct assurance testing.

Quality Assurance

North Dakota DOT is responsible for all assurance testing and acceptance. The engineer may select any or all of the split samples for assurance testing. Assurance testing is conducted on the properties evaluated for QC.

Acceptance

Acceptance is based on the contractor's QC tests if the results are verified by North Dakota DOT's assurance testing. The testing frequencies vary, depending on the property being measured. For density, the lot size is 1 day's production, with two core samples for each equal sublot being averaged.

Pay Factors

Pay factors are assigned for gradation, asphalt content, and density. The determination of the pay factors was not available; however, the total pay adjustment is determined by successively multiplying the pay factors for gradation, asphalt content, and density by the contract unit price.

QC/QA Tests for North Dakota

Table 89. North Dakota QC/QA tests.

Test	Sampling Location	QC Testing by Contractor	Assurance Testing Frequency by North Dakota DOT	Acceptance by North Dakota DOT
Gradation	Cold feed	One per 1500 tons	Engineer's discretion, minimum of 10% QC testing	Based on QC tests if verified by assurance test
Plastic index	Stockpiles	Three per 10,000 tons	Engineer's discretion, minimum of 10% QC testing	Based on QC tests if verified by assurance test
Lightweight aggregate	Stockpiles	Three per 10,000 tons	Engineer's discretion, minimum of 10% QC testing	Based on QC tests if verified by assurance test
Fractured faces	Stockpiles	Three per 10,000 tons	Engineer's discretion, minimum of 10% QC testing	Based on QC tests if verified by assurance test
MSG	Plant mix	One per 1500 tons	Engineer's discretion, minimum of 10% QC testing	Based on QC tests if verified by assurance test
BSG	Plant mix	One per 1500 tons	Engineer's discretion, minimum of 10% QC testing	Based on QC tests if verified by assurance test
Air voids	Plant mix or behind paver	One per 1500 tons	Engineer's discretion, minimum of 10% QC testing	Based on QC tests if verified by assurance test
Asphalt content	Plant mix or behind paver	Four per 1 day's production	Engineer's discretion, minimum of 10% QC testing	Based on QC tests if verified by assurance test
Density	Roadway	Two per sublot per 1 day's production	Engineer's discretion, minimum of 10% QC testing	Based on QC tests if verified by assurance test

1 ton = 0.907 Mg

Ohio

Quality Control

The contractor is responsible for developing the mix design for the project, providing a QC plan that must be approved by the engineer prior to construction, and providing qualified personnel and equipment to conduct QC testing. The contractor must allow the engineer access to the laboratory to observe any and all testing procedures, calculations, test documentation, and plotting of results.

The contractor is responsible for QC for the following properties: gradation, asphalt content, air voids, MSG, and density. The contractor must provide split samples to the engineer for assurance testing.

Quality Assurance

Ohio DOT is responsible for all assurance testing and acceptance of the material. Assurance testing is conducted on the properties evaluated for QC.

Acceptance

The contractor's control test can be used for acceptance for gradation, density, and asphalt content if the results are verified by Ohio DOT's assurance tests.

Pay Factors

A pay factor is assigned for gradation, density, and asphalt content. The determination of the pay factors was not provided. However, the pay factors are cumulative. The sum is then multiplied by the contract unit price to determine the amount of the bonus or penalty.

QC/QA Tests for Ohio

Table 90. Ohio QC/QA tests.

Test	Sample Location	QC Testing Frequency by Contractor	Assurance Testing Frequency by Ohio DOT	Acceptance by Ohio DOT
Asphalt content	Behind paver	Two per day or one for each 1400 tons	One per 2800 tons	Based on QC tests as verified by QA tests
Gradation (dry)	Cold feed	Two per day or one for each 1400 tons	One per 2800 tons	Based on QC tests as verified by QA tests
Air voids	Behind paver	Two per day or one for each 1400 tons	One per 2800 tons	NA
MSG	Behind paver	Two per day or one for each 1400 tons	One per 2800 tons	NA
Density	Roadway	Two per sublot (700 tons)	One per 2800 tons	Based on QC tests as verified by QA tests

NA = Not applicable
1 ton = 0.907 Mg

Ontario

Quality Control

The contractor is responsible for developing the mix design for the project, providing a QC plan that must be approved by the contract administrator prior to construction, and providing qualified personnel and equipment to conduct QC testing. The contractor must allow assurance samples to be taken to a certified laboratory at the request of the owner.

The contractor is responsible for QC for the following properties: gradation, asphalt content, air voids, and density.

Quality Assurance

The owner is responsible for all assurance testing. Owner assurance testing is conducted on gradation, air voids, VMA, Marshall stability, Marshall flow, and asphalt content.

Acceptance

The contractor is responsible for all acceptance testing. The contractor's control test can be used for acceptance if the results agree with the owner's assurance results. In the event that the two results do not agree, additional samples will be taken for referee testing by an independent laboratory. Acceptance testing is conducted for the following characteristics: gradation, air voids, VMA, Marshall stability, Marshall flow, and asphalt content. Acceptance is based on the PWL for each lot. Sublot sizes vary and are decided on by the owner and the contractor, but usually consist of 500-metric ton sublots.

Pay Factors

Pay factors are assigned for gradation, in-place density, asphalt content, and air voids. The individual pay factors are weighted and summed to arrive at the composite pay factor. The composite pay factor is then multiplied by the contract unit price to determine the amount of the bonus or penalty.

QC/QA Tests for Ontario

Table 91. Ontario QC/QA tests.

Test	Sampling Location	QC Testing Frequency by Contractor	Assurance Testing Frequency by Owner	Acceptance Testing Frequency by Contractor
Gradation	Cold feed	As needed for control	One per 500 metric tons	One per 500 metric tons
Asphalt content	Behind paver	As needed for control	One per 500 metric tons	One per 500 metric tons
Air voids	Behind paver	As needed for control	One per 500 metric tons	One per 500 metric tons
VMA	Behind paver	As needed for control	One per 500 metric tons	One per 500 metric tons
Marshall stability	Behind paver	As needed for control	One per 500 metric tons	One per 500 metric tons
Marshall flow	Behind paver	As needed for control	One per 500 metric tons	One per 500 metric tons
Density	Roadway	As needed for control	Three per sublot	Three per sublot

Oregon

Quality Control

The contractor is responsible for developing the mix design for the project, providing a QC plan that must be approved by the engineer prior to construction, and providing qualified personnel and equipment to conduct QC testing. The contractor must allow the engineer access to the laboratory to observe any and all testing procedures, calculations, test documentation, and plotting of results.

The contractor is responsible for QC for the following properties: gradation, asphalt content, MSG, moisture content, air voids, VMA, TSR, dust-to-asphalt ratio, and density. The contractor must supply the engineer with split samples for assurance testing.

Quality Assurance

Oregon DOT is responsible for all assurance testing. Assurance testing is conducted on the properties evaluated for QC. All control and assurance tests are conducted on 1000-Mg sublots.

Acceptance

The contractor's control test cannot be used for acceptance. Acceptance testing is conducted for the following characteristics: gradation, in-place density, asphalt content, and moisture content. Acceptance is on a lot-by-lot basis where a lot is represented by all the material produced under the JMF. Test results are used to establish an acceptable quality level using quality-level analysis.

Pay Factors

Pay factors are assigned for gradation, in-place density, asphalt content, and moisture. Once the PWL for an element is determined, the pay factor is taken from a table and then weighted according to the following equation:

$$WPF = (PF) \times f_i \tag{48}$$

where: f_i = weighting factor

The following table determines these factors:

Weighting Factors

Table 92. Oregon weighting factors.

Constituent	Weighting *f* Factor
Aggregate passing 37.5-, 31.5-, 25.0-, 19.0-, and 12.5-mm sieves	1 each
Aggregate passing 6.3-mm sieve	5
Aggregate passing 2.00-mm sieve	5
Aggregate passing 425-µm sieve	3
Aggregate passing 75-µm sieve	10
Asphalt content	26
Moisture content	8
Density	40

The composite pay factor is then determined by the following equation:

$$CPF = \sum WPF / \sum f_i \tag{49}$$

The composite pay factor is then multiplied by the quantity in the lot and by the contract unit price to determine the total amount of payment for the lot. The following table is a summary of the QC/QA tests:

QC/QA Tests for Oregon

Table 93. Oregon QC/QA tests.

Test	Sampling Location	QC Testing by Contractor	Assurance Testing Frequency by Oregon DOT	Acceptance Testing Frequency by Oregon DOT
Gradation	Cold feed	One per 1000 Mg	One per 1000 Mg	One per lot
Dust-to-asphalt ratio	Cold feed	One per 1000 Mg	One per 1000 Mg	NA
Asphalt content	Truck or behind paver	One per 1000 Mg	One per 1000 Mg	One per lot
Moisture	Truck	One per 1000 Mg	One per 1000 Mg	One per lot
Air voids	Truck or behind paver	One per 1000 Mg	One per 1000 Mg	NA
VMA	Truck or behind paver	One per 1000 Mg	One per 1000 Mg	NA
TSR	Truck or behind paver	One per 1000 Mg	One per 1000 Mg	NA
MSG	Truck or behind paver	One per 1000 Mg	One per 1000 Mg	NA
Density	Roadway	One per 1000 Mg	One per 1000 Mg	Five per lot

NA = Not applicable

Pennsylvania

Quality Control

The contractor is responsible for developing the mix design for the project, providing a QC plan that must be approved by the engineer prior to construction, and providing qualified personnel and equipment to conduct QC testing. The contractor must allow the engineer access to the laboratory to observe any and all testing procedures, calculations, test documentation, and plotting of results.

The contractor is responsible for QC for the following properties: gradation, stability, flow, voids, asphalt content, mix temperature, and density. These properties must be tested once per 250-ton (230 Mg) sublot.

Quality Assurance

Pennsylvania DOT is responsible for all assurance sampling and testing. The assurance tests are independent of the contractor's control tests. Assurance testing is conducted for gradation, asphalt content, and density.

Acceptance

Acceptance is based on certification at the plant when QC and assurance tests conform to the JMF tolerances, except for in-place density. Acceptance for density is accomplished by verification of the QC tests by the QA tests in the field. Acceptance properties are asphalt content, gradation (#200), stability, flow voids, and density.

Pay Factors

Pay factors are assigned for gradation, in-place density, and asphalt content. The pay factors are based on the deviation from the specification limits. The following tables provide the pay factor for asphalt content, gradation, and density:

Adjustment of Contract Price Relative to the Specification Limits

Table 94. Pennsylvania adjustment of contract price relative to the specification limits.

	Test Value	Payment Factor Percentage
Asphalt Content	±0.07%	100
	±0.8-1.0%	75
	> ±1.0%	*
Percentage Passing #200 Sieve	±3.0%	100
	±3.1-4.0%	75
	> ±4.0%	*
Density	≥ 92% or < 97% of TMD	100
	90-91% or 97-99% of TMD	98
	≤ 89% or > 99% of TMD	*

*Indicates remove and replace.

QC/QA Tests for Pennsylvania

Table 95. Pennsylvania QC/QA tests.

Test	Sampling Location	QC Testing Frequency by Contractor	Assurance Testing Frequency by Pennsylvania DOT	Acceptance by Pennsylvania DOT
Gradation (dry)	Cold feed	One per sublot (250 tons)	One per lot (1000 tons)	Based on plant certification by QC and assurance tests
Mixture temperature	Truck	One per sublot (250 tons)	NA	NA
Flow	Plant	One per sublot (250 tons)	NA	Based on plant certification by QC and assurance tests
Air voids	Plant	One per sublot (250 tons)	NA	Based on plant certification by QC and assurance tests
Asphalt content	Plant	One per sublot (250 tons)	One per lot (1000 tons)	Based on plant certification by QC and assurance tests
Density	Roadway	One per sublot (250 tons)	One per lot (1000 tons)	Based on QC as verified by QA
VMA	Plant	One per sublot (250 tons)	NA	Based on plant certification by QC and assurance tests

NA = Not applicable
1 ton = 0.907 Mg

South Carolina

Quality Control

The contractor is responsible for all QC testing. Prior to the start of construction, the contractor must submit a QC program for approval by the engineer. This plan must include, at a minimum, the name of the QC manager, sampling frequency and procedure, testing procedures, method of recording results, and the procedure for managing variable test results. The contractor must provide a certified level 1 technician to monitor gradation and binder content. The contractor is also required to develop and submit a JMF that conforms to all South Carolina DOT requirements. The technician must make all test results available to the engineer and the engineer must be permitted to observe all testing of material.

The JMF is verified based on gradation, asphalt content, Marshall stability, Marshall flow, air voids, TSR, VMA, and dust-to-asphalt ratio.

The contractor shall use the test methods identified in the table below to perform the plant QC tests and verifications at a frequency not less than that indicated.

Required QC Tests and Verifications

Table 96. South Carolina required QC tests and verifications.

Parameter	Minimum Frequency	Sampling Method	Test Method
Bituminous mixture gradation	One per lot	SC-T-62	SC-T-76
MSG (Rice method)	Two per lot	SC-T-62	SC-T-83
Marshall stability	One per lot	SC-T-62	SC-T-66
Lime rate verification	Two per lot	SC-T-71	SC-T-71 SC-T-78
Mixture temperature verification	Four per lot	—	SC-T-84
Aggregate stockpile gradation	One per 10,000 tons	SC-T-1, SC-T-2	SC-T-4
Temperature Ambient air	Before paving starts, then two per lot	—	SC-T-84
Mat	Four per lot	—	SC-T-84
Calculated lay-down rate	One per 200 tons	—	SC-T-85
Tack rate, type, and dilution	One per application	—	SC-T-86

1 ton = 0.907 Mg

Compaction control and the number and types of rollers shall be the contractor's responsibility. The contractor shall have a nuclear density gauge at the site during all HMA placement and compaction operations and shall use the gauge to assist in QC of the compaction process.

Quality Assurance

The contractor is responsible for conducting all acceptance testing. South Carolina DOT personnel may witness the sampling and testing being performed by the contractor. South Carolina DOT will conduct its own tests to verify the contractor's test results. The verification tests for asphalt binder content will be on the split samples retained by the contractor. For most paving, the verification tests for in-place density will be on cores retained by the contractor. For thin-lift surface courses, verification tests for in-place density analysis will be at independent locations determined and tested by South Carolina DOT using the same type of nuclear density gauge as used by the contractor.

In general, the frequency of South Carolina DOT's verification tests will be equal to or greater than 10 percent of the tests required for the contractor.

Acceptance

The contractor shall use the test methods identified in the table below to perform the plant acceptance tests at a frequency not less than that indicated.

Required Acceptance Tests

Table 97. South Carolina required acceptance tests.

Test Parameter	Typical Frequency	Sampling Method	Test Method
Asphalt binder content, percent	Four per lot	SC-T-100 SC-T-62	SC-T-64 or SC-T-75
Voids analysis Air voids, percent VMA, percent	 Four per lot Four per lot	 SC-T-100, SC-T-62 SC-T-100, SC-T-62	 SC-T-68 SC-T-68
In-place density (percentage of maximum theoretical) *Note: Requirements apply to intermediate courses and to surface courses other than type 3 and type 4.*	One per 1500-ft sublot	SC-T-100, SC-T-87	SC-T-87
In-place density (percentage of target nuclear control-strip density) *Note: Requirements apply to type 3 and type 4 surface courses and for all thin-lift surface courses regardless of type.*	10 per lot	SC-T-100	SC-T-65

1 ft = 0.305 m

The engineer is responsible for determining the air voids and VMA from Marshall samples. The technician is responsible for determining the binder content. Density is determined by coring and is the responsibility of the contractor at the direction of the engineer. Lot size is 1 day's production. Test frequency is at the discretion of the engineer; however, the minimum requirements are listed in the table.

Pay Factors

For asphalt binder content and the Marshall volumetric properties of air voids and VMA, a *lot* is normally defined as 1 day's production. The acceptance and pay factors for asphalt binder content and the Marshall volumetric properties of air voids and VMA will be based on the percentage of the lot that is within the specification limits (PWL) based on the quality index calculated using the test results from the lot.

The pay factor for each property is determined based on the total percent within limits (*TPWL*) value from the following equation:

$$PF = 55 + 0.5(TPWL) \tag{50}$$

The percent pay factor for the lot (*LPF*) will be determined by multiplying the percent pay factors for asphalt binder content, air voids, VMA, and in-place density by weighted coefficients as shown in the equation below. The *LPF* will be calculated to the nearest tenth (0.1) and rounded up to the next whole percent (1.0). The *LPF* will be determined from the following equation:

$$LPF = 0.2(PF_{AC}) + 0.35(PF_{AV}) + 0.1(PF_{VMA}) + 0.35(PF_{Den}) \quad (51)$$

where:
- LPF = percent pay factor for the lot
- PF_{AC} = percent pay factor for asphalt binder content
- PF_{AV} = percent pay factor for air voids
- PF_{VMA} = percent pay factor for VMA
- PF_{Den} = percent pay factor for in-place density

Texas

Quality Control
The contractor is responsible for providing a QC plan that must be approved by the engineer prior to construction and providing qualified personnel and equipment to conduct QC testing. The contractor must allow the engineer access to the laboratory to observe any and all testing procedures, calculations, test documentation, and plotting of results. To complete the required testing, the contractor must provide technicians certified at levels I through IV, depending on the test being conducted. The standard lot size is 6000 square yards (5000 m^2) of surface area of PCC pavement of the same thickness. A sublot is equal to one-fifth of the surface area of a lot. If the final lot is 3000 square yards (2500 m^2) or greater, it will be considered a full lot. QC must be exercised on all properties contained in the table entitled QC/QA Tests for Texas.

Quality Assurance
QA is accomplished through verification testing conducted by Texas DOT. Independent assurance testing is conducted to review the QC testing and to check the accuracy of the equipment used for verification testing.

Acceptance
Acceptance of the material is based on the QC tests and the verification tests. If the contractor's QC test results and the engineer's verification test results are not within tolerances, referee testing will be conducted. If referee tests are requested, the referee test results shall govern.

Pay Factors
Pay factors are assigned for ride quality, thickness, and flexural strength. The pay factor for flexural strength is based on the following table:

Table 98. Texas pay factors for flexural strength.

Average Flexural Strength (Three Beams) (lbf/inch^2)	Strength Adjustment Factor (SF)
≥ 555	1.0000
550	0.9695
545	0.9397
540	0.9106
535	0.8820
530	0.8542
525	0.8269
< 525	Remove or remain with no payment

1 pound force per square inch (lbf/inch^2) = 6.89 kilopascals (kPa)

The pay factor for ride quality was not provided. The pay factor for thickness is determined using the following table:

Table 99. Texas pay factors for thickness.

Thickness Deficiency (inch)	Thickness Adjustment Factor (TF)
> 0.00-0.20	1.00
> 0.20-0.30	0.80
> 0.30-0.40	0.72
> 0.40-0.50	0.68
> 0.50-0.75	0.57

1 inch = 25.4 mm

The pay adjustment factor for each lot is calculated as AF = (SF)(TF). The pay adjustment for each lot is then calculated to be Payment = (BP)(AF)(Q), where *BP* = bid price and *Q* = lot quantity of acceptable pavement.

QC/QA Tests for Texas

Table 100. Texas QC/QA tests.

Test	Sample Location	QC Testing by Contractor	Verification Testing by Texas DOT	Acceptance by Texas DOT
Mix temperature	Plant	One per sublot	One per day	Based on QC, verification, and referee, if required
Air content (plastic)	Roadway after discharge	One per sublot	1 per 10 QC tests	Based on QC, verification, and referee, if required
Concrete unit weight	Roadway after discharge	One per sublot	1 per 10 QC tests	Based on QC, verification, and referee, if required
Making/curing strength specimens	Roadway after discharge	One per sublot	1 set per 10 sublots	NA
7-day flexural strength	Roadway after discharge	NA	One per sublot	Based on verification
Coarse aggregate gradation	Stockpile	One per stockpile per day	1 per 10 QC tests	Based on QC, verification, and referee, if required
Coarse aggregate loss by decantation	Stockpile	One per 5 PCC production days	1 per 10 QC tests	Based on QC, verification, and referee, if required
Fine aggregate gradation	Stockpile	One per stockpile per day	1 per 10 QC tests	Based on QC, verification, and referee, if required
Fineness modulus	Stockpile	One per stockpile per day	1 per 10 QC tests	Based on QC, verification, and referee, if required
Fine aggregate organic impurity	Stockpile	One per 5 PCC production days	1 per 10 QC tests	Based on QC, verification, and referee, if required
Sand equivalent	Stockpile	One per 5 PCC production days	1 per 10 QC tests	Based on QC, verification, and referee, if required
Aggregate moisture content	Stockpile	One per stockpile per day	1 per 10 QC tests	Based on QC, verification, and referee, if required
Water-to-cement ratio	Roadway after discharge	One per batch	Review plant printout and calibration	Based on QC, verification, and referee, if required
Cement factor	Roadway after discharge	One per batch	Review plant printout and calibration	Based on QC, verification, and referee, if required
Admixture dosage	Roadway after discharge	One per batch	Review plant printout and calibration	Based on QC, verification, and referee, if required
Coring for thickness	Roadway	One per sublot	NA	NA
Measuring pavement thickness	Roadway (plastic)	NA	One per sublot	Based on QC, verification, and referee, if required

Virginia

Quality Control

The contractor is responsible for developing the mix design for the project, providing a QC plan that must be approved by the engineer prior to construction, and providing qualified personnel and equipment to conduct QC testing. The contractor must allow the engineer access to the laboratory to observe any and all testing procedures, calculations, test documentation, and plotting of results.

Mix design verification is evaluated for the following properties: Marshall stability, Marshall flow, MSG, gradation, BSG, VMA, VFA, VTM, dust-to-asphalt ratio, and asphalt viscosity.

The contractor is responsible for QC for the following properties: air voids, VFA, VMA, dust-to-asphalt ratio, density, temperature, and fractured faces. The contractor must cooperate with the engineer to obtain samples for assurance testing. The contractor must perform at least one set of tests per day or one set per 1000 metric tons.

Quality Assurance

QA is achieved through the QC testing by the contractor and by independent monitoring (assurance) tests conducted by Virginia DOT. Assurance testing may be conducted on air voids, VFA, VMA, dust-to-asphalt ratio, density, temperature, and fractured faces. The frequency of testing is at the engineer's discretion.

Acceptance

The contractor's control test can be used for acceptance if the results are verified by the assurance tests conducted by Virginia DOT. The testing frequencies vary, depending on the property being measured.

Pay Factors

Pay factors are assigned for gradation and asphalt content. The pay factors are assigned on an adjustment point system. Adjustment points are assigned for each sieve size and for asphalt content. Any lot receiving 25 or more points must be removed. Any lot with fewer than 25 adjustment points may be left in place with a 1-percent reduction in the unit bid price per point. A summary of the QC/QA tests follows:

QC/QA Tests for Virginia

Table 101. Virginia QC/QA tests.

Test	Sampling Location	QC Testing by Contractor	Assurance Testing Frequency by Virginia DOT	Acceptance by Virginia DOT
Gradation (extracted)	Truck	One per 500 tons or four per lot (2000 tons)	At the discretion of the engineer for QC verification	Based on observed and verified results of QC by assurance tests
Asphalt content	Truck	One per 500 tons or four per lot (2000 tons)	At the discretion of the engineer for QC verification	Based on observed and verified results of QC by assurance tests
Air voids	Truck	One Marshall test per day or 1000 tons	At the discretion of the engineer for QC verification	Based on observed and verified results of QC by assurance tests
VFA	Truck	One Marshall test per day or 1000 tons	At the discretion of the engineer for QC verification	Based on observed and verified results of QC by assurance tests
VMA	Truck	One Marshall test per day or 1000 tons	At the discretion of the engineer for QC verification	Based on observed and verified results of QC by assurance tests
Fractured faces	Cold feed	One Marshall test per day or 1000 tons	At the discretion of the engineer for QC verification	Based on observed and verified results of QC by assurance tests
Dust-to-asphalt ratio	Cold feed	One Marshall test per day or 1000 tons	At the discretion of the engineer for QC verification	Based on observed and verified results of QC by assurance tests
Density	Roadway	Based on nuclear gauge testing of control strip	At the discretion of the engineer for QC verification	Based on nuclear gauge testing of control strip
VTM	Truck	One Marshall test per day or 1000 tons	At the discretion of the engineer for QC verification	Based on observed and verified results of QC by assurance tests

1 ton = 0.907 Mg

Washington

Quality Control

The contractor must submit materials to be used in the mix to Washington State DOT for mix design development. The contractor must allow 15 working days for the development and approval of the design. In addition, the contractor is responsible for providing a QC plan that must be approved by the engineer prior to construction and providing qualified personnel and equipment to conduct QC testing. The contractor must allow the engineer access to the laboratory to observe any and all testing procedures, calculations, test documentation, and plotting of results.

The contractor is responsible for QC for the following properties: gradation, asphalt content, density, fractured faces, and sand equivalent.

The frequency of the QC testing is established and agreed upon by the contractor and the engineer in a preconstruction conference establishing the control plan.

Quality Assurance

Washington State DOT is responsible for all assurance sampling and testing. The assurance testing frequencies vary, depending on the property being measured and are established by the engineer. Assurance testing is conducted on the properties evaluated for QC.

Acceptance

Acceptance testing is the responsibility of Washington State DOT. The contractor's control test cannot be used for acceptance. Acceptance is based on the following characteristics: gradation, in-place density, smoothness, and asphalt content. A lot is considered to be all material produced using the same JMF. Sublot sizes vary, but cannot exceed 800 tons (730 Mg). Acceptance and the corresponding pay factors are based on the PWL.

Pay Factors

Pay factors are assigned for gradation, in-place density, and asphalt content. The individual pay factors for an element are determined using quality-level analysis. These individual pay factors are weighted with an *f*-factor. The composite pay factor is then determined according to the following equation:

$$CPF = f_1(PF_1) + f_2(PF_2) + \ldots f_i(Pf_i) / \sum f_i \qquad (52)$$

QC/QA Tests for Washington

Table 102. Washington QC/QA tests.

Test	Sample Location	QC Testing Frequency by Contractor	Assurance Testing Frequency by Washington State DOT	Acceptance Testing Frequency by Washington State DOT
Extracted gradation (⅝ inch, ½ inch, ⅜ inch, ¼ inch, #10, #40, #200)	Truck	Engineer's discretion	Engineer's discretion	Five per lot minimum
Asphalt content	Truck	Engineer's discretion	Engineer's discretion	Five per lot minimum
Density	Roadway	Engineer's discretion	Engineer's discretion	Five per lot minimum
Smoothness	Roadway	Engineer's discretion	Engineer's discretion	Based on one lot
Fractured faces	Cold feed	Engineer's discretion	Engineer's discretion	Five per lot minimum

1 inch = 25.4 mm

Wisconsin

Quality Control
The contractor is responsible for developing the mix design for the project, providing a QC plan that must be approved by the engineer prior to construction, and providing qualified personnel and equipment to conduct QC testing. The contractor must allow the engineer access to the laboratory to observe any and all testing procedures, calculations, test documentation, and plotting of results.

The contractor must conduct QC tests for the following properties: gradation (1/2 inch (13 mm), 3/8 inch (10 mm), #4, #8, #30, and #200), asphalt content, Marshall BSG, MSG, air voids, and VMA.

The contractor must supply the engineer with split samples for assurance testing. The number of samples required depends on the amount of material produced. The number of required tests is as follows:

Sampling Frequencies

Table 103. Wisconsin sampling frequencies.

Total Daily Plant Production (Mg)	Number of Samples per Day
45-550	1
551-1360	2
1361-2450	3
2451-3810	4
3811+	Add one sample for each additional 1360 Mg or part thereof

Quality Assurance
Wisconsin DOT is responsible for all assurance testing and acceptance. The assurance testing frequencies must be at least 10 percent of the contractor's QC frequency. Assurance testing is conducted on the properties evaluated for QC.

Acceptance
The contractor's control tests are used for acceptance if the results are verified by Wisconsin DOT's assurance tests.

Pay Factors
Pay factors are assigned for gradation, asphalt content, air voids, and VMA. The pay factors are assigned if the contractor's test results and the assurance test results are not within specified tolerances. The pay factors are presented in the following table:

Percent Payment for Mixture

Table 104. Wisconsin percent payment for mixture.

Item	Produced Within Warning Bands	Produced Outside JMF Limits
Gradation	90	75
Asphalt content	85	75
Air voids	70	50
VMA	90	75

The minimum single payment applies.

QC/QA Tests for Wisconsin

Table 105. Wisconsin QC/QA tests.

Test	Sampling Location	QC Testing by Contractor	Assurance Testing Frequency by Wisconsin DOT	Acceptance by Wisconsin DOT
Gradation (extracted)	Truck	See Sampling Frequency table	10% of QC frequency	Based on verification of QC tests by assurance tests
Asphalt content	Truck	See Sampling Frequency table	10% of QC frequency	Based on verification of QC tests by assurance tests
BSG	Truck	See Sampling Frequency table	10% of QC frequency	Based on verification of QC tests by assurance tests
MSG	Truck	See Sampling Frequency table	10% of QC frequency	Based on verification of QC tests by assurance tests
Air voids	Truck	See Sampling Frequency table	10% of QC frequency	Based on verification of QC tests by assurance tests
VMA	Truck	See Sampling Frequency table	10% of QC frequency	Based on verification of QC tests by assurance tests

Wyoming

Quality Control

The contractor is responsible for developing the mix design for the project, providing a QC plan that must be approved by the engineer prior to construction, and providing qualified personnel and equipment to conduct QC testing. The contractor must allow the engineer access to the laboratory to observe any and all testing procedures, calculations, test documentation, and plotting of results.

The contractor is responsible for QC for the following properties: gradation, liquid limit and plastic index of virgin material, moisture content of mix, in-place density, test strip, and mix verification.

Mix design verification is evaluated for the following properties: LA abrasion, number of Marshall blows, Marshall stability, Marshall flow, percent voids in mix, dust-to-asphalt ratio, minimum asphalt percentage, minimum TSR, film thickness, and VMA.

Quality Assurance

Wyoming DOT is responsible for all assurance testing. Assurance testing is conducted on the properties evaluated for QC.

Acceptance

The contractor's control test cannot be used for acceptance. The testing frequencies vary, depending on the property being measured. Acceptance testing is conducted on the following characteristics: gradation, in-place density, and asphalt content.

Pay Factors

Pay factors are assigned for gradation, in-place density, and asphalt content. Determination of the pay factors uses quality-level analysis. Once the quality level and corresponding pay factor are determined, the pay adjustments are calculated using the following equations:

$$PA_A = 0.67 \times PMP \times (PF_A - 1) \times LS_A \tag{53}$$

$$PA_D = 1.33 \times PMP \times (PF_D - 1) \times LS_D \tag{54}$$

$$PA_{AC} = 0.67 \times PMP \times (PF_{AC} - 1) \times LS_{AC} \tag{55}$$

where:
- PA_A = pay adjustment for aggregate gradation
- PF_A = pay factor for gradation
- PA_D = pay adjustment for density
- PF_D = pay factor for density
- PA_{AC} = pay adjustment for asphalt content
- PF_{AC} = pay factor for asphalt
- LS_A = lot size for aggregate gradation evaluation
- LS_D = lot size for density evaluation
- LS_{AC} = lot size for asphalt content evaluation
- PMP = bituminous mixture unit price

QC/QA Tests for Wyoming

Table 106. Wyoming QC/QA tests.

Test	Sample Location	QC Testing by Contractor	Assurance Testing Frequency by Wyoming DOT	Acceptance Testing Frequency by Wyoming DOT
Stockpile gradation	Stockpile	One per 1000 tons	One per 1000 tons	One per lot or 5000 tons
Liquid limit and plastic index (virgin aggregate)	Individual stockpiles	One per 1000 tons	One per 1000 tons	NA
Virgin aggregate gradation	Cold feed	One per 1000 tons	One per 1000 tons	One per lot or 5000 tons
Mix verification	Truck	One per day for 4 days, then one per 20,000 tons	One per day for 4 days, then one per 20,000 tons	NA
Moisture content of mix	Truck	One per day	One per day	NA
Test strip	Roadway	Required	NA	NA
In-place density	Roadway	One per 200 tons	One per 200 tons	Seven per lot or 1500 tons
Asphalt content	Truck	One per day	One per day	One per day

NA = Not applicable
1 ton = 0.907 Mg

APPENDIX C: SUMMARY OF AGENCY SUPERPAVE SPECIFICATIONS RECEIVED

Connecticut

Quality Control

The contractor is responsible for developing the mix design for the project, providing a QC plan that must be approved by the engineer prior to construction, and providing qualified personnel and equipment to conduct QC testing. The contractor must allow the engineer access to the laboratory to observe any and all testing procedures, calculations, test documentation, and plotting of results.

Mix design verification is evaluated for the following properties: gradation, asphalt content, voids @ N_{des}, VMA, VFA, G_{mm} @ N_{des} and N_{max}, G_{mb}, G_{se}, TSR, N_{des}, N_{ini}, N_{max}, recycled asphalt pavement (RAP) gradation, fine aggregate angularity (FAA), coarse aggregate angularity (CAA), sand equivalent, and dust-to-asphalt ratio.

The contractor must exercise control over the following properties: asphalt content, gradation, air voids, mix moisture, VMA, VFA, BSG of mix, MSG of mix, RAP gradation, aggregate moisture, RAP moisture, and density. In addition, the contractor must provide split samples to the engineer for verification testing.

Quality Assurance

Connecticut DOT is responsible for QA. Assurance is accomplished through verification testing by Connecticut DOT and through independent assurance testing. Verification tests are conducted on split samples. A lot is considered 1 day’s production. Sublots are five equal divisions of any lot.

Acceptance

Acceptance is based on the QC tests as verified by Connecticut DOT’s verification tests. Density acceptance is based on the PWL of a lot. The PWL is calculated using 10 tests (2 per sublot).

Pay Factors

Pay factors are assigned for gradation, asphalt content, joint density, and mat density. The pay factors for each can be determined from the following tables. Total payment is determined by successively multiplying the percent payment by the contract price per metric ton of material.

Pay Factors for Gradation and Binder Content

Maximum Allowable Deviation (M.A.D.) From Job-Mix Formula for Consecutive Tests

Table 107. Connecticut M.A.D. from job-mix formula for consecutive tests.

Item	M.A.D.	Percent Payment
Binder	±0.4	90
75	±2	90
150	±3	90
300	±3	90
600	±4	90
1.18	±4	90
2.36	±6	90
4.75	±6	90
9.5	±6	90
12.5	±6	90
19.0	±6	90
25.0	±6	90
37.5	±6	90
50.0	±6	90

Pay Factors for Joint and Mat Density

Table 108. Connecticut pay factors for joint and mat density.

PWL	Percent Payment Mat Density	Percent Payment Joint Density
90-100	100	100
80-90	$0.5PWL + 55$	$0.5PWL + 55$
65-80	$2.0PWL - 65$	$2.0PWL - 65$
< 65	0 or remove	50

QC/QA Tests for Connecticut

Table 109. Connecticut QC/QA tests.

Test	Sampling Location	QC Testing Frequency by Contractor	Verification Testing Frequency by Connecticut DOT	Acceptance by Connecticut DOT based on …
Gradation	Conveyor prior to binder addition	Five per lot	One per lot	QC verified by Connecticut DOT's test
Moisture content of RAP	Stockpile	Two per day	One per lot	QC verified by Connecticut DOT's test
Moisture content of aggregate	Conveyor prior to binder addition	Two per day	One per lot	QC verified by Connecticut DOT's test
Asphalt content of RAP	Stockpile	Two per day	One per lot	QC verified by Connecticut DOT's test
Rap gradation	Stockpile	Two per day	One per lot	QC verified by Connecticut DOT's test
Asphalt content	Behind paver	Five per lot	One per lot	QC verified by Connecticut DOT's test
Superpave molds @ N_{max}	Behind paver	Five sets (two molds) per lot	One per lot	QC verified by Connecticut DOT's test
BSG	Behind paver	Five per lot	One per lot	QC verified by Connecticut DOT's test
Air voids	Behind paver	Five per lot	One per lot	QC verified by Connecticut DOT's test
VMA	Behind paver	Five per lot	One per lot	QC verified by Connecticut DOT's test
VFA	Behind paver	Five per lot	One per lot	QC verified by Connecticut DOT's test
MSG of mix	Plant	Five per lot	One per lot	QC verified by Connecticut DOT's test
Mat density (nuclear)	Roadway	10 per lot	NA	Based on PWL for lot
Joint density (nuclear)	Roadway	10 per lot	NA	Based on PWL for lot

NA = Not applicable

Kansas

Quality Control

The contractor is responsible for developing the mix design for the project, providing a QC plan that must be approved by the engineer prior to construction, and providing qualified personnel and equipment to conduct QC testing. All testing equipment and procedures must conform to the Kansas DOT *Superpave Mix Design and Superpave Field Laboratory Technician Certification Training Manual*. The contractor must allow the engineer access to the laboratory to observe any and all testing procedures, calculations, test documentation, and plotting of results.

Mix design verification is evaluated for the following properties: gradation, asphalt content, theoretical MSG, BSG, air voids, VMA, VFA, dust-to-asphalt ratio, flat or elongated particles, percent moisture, and TSR.

The contractor must exercise control over the following properties: asphalt content, gradation, air voids, VMA, VFA, CAA, sand equivalent, FAA, TSR, density, and dust-to-asphalt ratio. Samples for mixture properties, aggregate gradation, and binder must be taken from behind the paver and transported to the field laboratory for testing.

Quality Assurance

Kansas DOT is responsible for all assurance and verification testing. Kansas DOT assurance tests are independent of the contractor's QC tests. The lot size for verification testing is 3000 tons (2720 Mg) with 750-ton (680 Mg) sublots.

Acceptance

The contractor's control tests are used for acceptance unless the results do not compare favorably with Kansas DOT's verification test results. In this case, Kansas DOT's tests are used for acceptance unless the contractor requests referee testing. The testing frequencies vary, depending on the property being measured.

Pay Factors

Pay factors are assigned for in-place density, air voids, and smoothness. Pay factors for density and air voids can be determined from the following tables. Information concerning methods for determining a smoothness pay factor was not included.

Pay Factors for Specified Density

Table 110. Kansas pay factors for specified density.

Percentage of MSG, Pay Factor A[b] (Average of 10 Density Tests[a])	Pay Factor A[b]		
	Mixes SM-2A, SR-2A, SM-2C, and SR-2C	Mix SM-1T	Mixes SM-1B and SR-1B
≥ 94.0			1.020
≥ 93.0	1.020	1.020	
93.0-93.9			A4
92.0-92.9	A1	A1	1.000
91.0-91.9	1.000	1.000	
90.0-90.9		1.000	
88.0-91.9			A5
87.0-90.9	A2		
86.0-89.9		A3	
< 88.0			(c)
< 87.0	(c)		
< 86.0		(c)	
Lowest Average of Any Sublot Within the Lot	**Pay Factor B[b]**		
≥ 92.0			1.020
≥ 91.0	1.020	1.020	
91.0-90.9			B4
90.0-90.9	B1	B1	1.000
89.0-89.9	1.000	1.000	
88.0-88.9		1.000	
86.0-89.9			B5
85.0-88.9	B2		
84.0-87.9		B3	
< 86.0			(c)
< 85.0	(c)		
< 84.0		(c)	

[a] For low daily production rates of less than 1000 tons (910 Mg) or when the engineer's verification tests are to be used for density pay determination, the lot sample size is as determined under subsection 603.05(b) for compaction testing.
[b] For shoulders less than 1.8 m wide and containing rumble strips, the density pay factors will not apply for the top lift. Density will be controlled by using an approved rolling procedure.
[c] Engineer will determine whether the material may remain in place. The pay factor for such material remaining in place is 0.700 for pay factor A and 0.800 for pay factor B.

Calculations for Pay Factors A1, A2, A3, A4, A5, B1, B2, B3, B4, and B5:

A1 = [100 + 2 (percentage of lot MSG – 92.0)] ÷ 100
A2 = [90 + 2.5 (percentage of lot MSG – 87.0)] ÷ 100
A3 = [90 + 2.5 (percentage of lot MSG – 86.0)] ÷ 100
A4 = [100 + 2 (percentage of lot MSG – 93.0)] ÷ 100
A5 = [90 + 2.5 (percentage of lot MSG – 88.0)] ÷ 100
B1 = [100 + 2 (percentage of sublot MSG – 90.0)] ÷ 100
B2 = [90 + 2.5 (percentage of sublot MSG – 85.0)] ÷ 100
B3 = [90 + 2.5 (percentage of sublot MSG – 84.0)] ÷ 100
B4 = [100 + 2 (percentage of sublot MSG – 91.0)] ÷ 100
B5 = [90 + 2.5 (percentage of sublot MSG – 86.0)] ÷ 100

Pay Factor Calculation: Density Pay Adjustment Factor (P_D)* = [(Pay factor A)(Pay factor B)] – 1.000 (56)

* P_D will be rounded to the nearest 0.001

Air Void Payment

Payment adjustment for air voids will be computed by multiplying the air voids pay adjustment factor (P_V) times the number of tons included in the lot times the bid price per ton. Calculate P_V to 0.001.

A lot will normally be comprised of the results of four contiguous individual air void tests performed on Superpave gyratory compacted samples of a given mix design. Lot size is defined in subsections 5.0(f) and 5.0(g) of this special provision. (Air void lots and density lots are normally of different sizes.)

The absolute value of the deviation from the target at N_{des} for each individual air void test in a lot will be computed as shown:

$$\text{Deviation From Target} = 4.0 - \text{Test value} \quad (57)$$

The average deviation (D) from the target will be computed as follows:

$$D = \text{Sum of deviations from target} \div \text{number of tests in lot} \quad (58)$$

The average air void deviation (D) will be used to select the air void pay factor from the appropriate table below. Calculate D to 0.01.

Pay Factor Table for Air Voids (Lot Size of Four Tests)

Table 111. Kansas pay factors for air voids (lot size of four tests).

Air Void Average Deviation	Pay Factor
$0.00 \le D_4 \le 0.35$	1.030
$0.36 \le D_4 \le 0.55$	$1.000 + 0.15(0.55 - D)$
$0.56 \le D_4 \le 1.05$	1.000
$1.06 \le D_4 \le 1.40$	$1.000 - 0.44(D - 1.05)$
$1.41 \le D_4$	(a)

[a] Engineer will determine whether the material may remain in place. The pay factor for such material remaining in place is 0.8000.

When the testing rate does not provide a complete four-test lot (e.g., at the end of a production run), the following applies: Combine the three tests into a lot and use the following table for D_3. When there are one or two tests remaining, combine them with the previous four tests to create a five- or six-test lot, respectively. The pay factors for D_5 and D_6 are also shown below.

Pay Factor Table for Air Voids (Lot Size of Three Tests)

Table 112. Kansas pay factors for air voids (lot size of three tests).

Air Void Average Deviation	Pay Factor
$0.00 \le D_3 \le 0.37$	1.030
$0.38 \le D_3 \le 0.58$	$1.000 + 0.14(0.58 - D)$
$0.59 \le D_3 \le 1.11$	1.000
$1.12 \le D_3 \le 1.48$	$1.000 - 0.35(D - 1.11)$
$1.49 \le D_3$	(a)

[a] Engineer will determine whether the material may remain in place. The pay factor for such material remaining in place is 0.8000.

Pay Factor Table for Air Voids (Lot Size of Five Tests)

Table 113. Kansas pay factors for air voids (lot size of five tests).

Air Void Average Deviation	Pay Factor
$0.00 \le D_5 \le 0.29$	1.030
$0.30 \le D_5 \le 0.53$	$1.000 + 0.125(0.53 - D)$
$0.54 \le D_5 \le 1.01$	1.000
$1.02 \le D_5 \le 1.35$	$1.000 - 0.41(D - 1.01)$
$1.36 \le D_5$	(a)

(a) Engineer will determine whether the material may remain in place. The pay factor for such material remaining in place is 0.8000.

Pay Factor Table for Air Voids (Lot Size of Six Tests)

Table 114. Kansas pay factors for air voids (lot size of six tests).

Air Void Average Deviation	Pay Factor
$0.00 \le D_6 \le 0.28$	1.030
$0.29 \le D_6 \le 0.51$	$1.000 + 0.13(0.51 - D)$
$0.52 \le D_6 \le 0.98$	1.000
$0.99 \le D_6 \le 1.31$	$1.000 - 0.39(D - 0.98)$
$1.32 \le D_6$	(a)

(a) Engineer will determine whether the material may remain in place. The pay factor for such material remaining in place is 0.8000.

Pay Factor Calculation: Air Void Pay Adjustment Factor (P_V) = Pay factor – 1.000 (59)

QC/QA Tests for Kansas

Table 115. Kansas QC/QA tests.

Test	Sampling Location	QC Testing Frequency by Contractor	Verification Testing Frequency by Kansas DOT	Assurance Testing Frequency by Kansas DOT
Binder sampling	Plant	One per three loads	One per project	NA
Binder content	Behind paver	One per sublot	One per three lots	Witness and test once per year by certified technician
Mix gradation (extraction)	Behind paver	One per sublot	One per three lots	Witness and test once per year by certified technician
Theoretical MSG	Behind paver	One per sublot	One per lot	Witness and test once per year by certified technician
Air voids	Behind paver	One per sublot	One per lot	Compact and test split sample once per week or 15,000 tons
RAP binder content	Stockpile	One per 1000 tons	One per 20,000 tons	NA
RAP gradation	Stockpile	One per 1000 tons	One per project	NA
Moisture damage to mix	Behind paver	One on first lot, then one per week or 10,000 tons	One per project	Witness and test once per year by certified technician
Sand equivalent	From conveyor	One per lot	One per project	Witness and test once per year by certified technician
CA angularity	Stockpile	One per 1000 tons	One per week or 10,000 tons	Witness and test once per year by certified technician
FA angularity	Stockpile	One on first lot, then one per 10,000 tons	One per project	Witness and test once per year by certified technician
Gradation of individual aggregate	Stockpile	One per 1000 tons	One per project per aggregate	Witness and test once per year by certified technician
Percent moisture in mix	Behind paver	One per lot	One per project	NA
Percent moisture in combined aggregate	From conveyor	One per lot	One per project	NA
Percent moisture in RAP	Stockpile	One per lot	One per project	NA
Plasticity index and gradation of mineral filler	Stockpile	One per 250 tons	One per project	Witness and test once per year by certified technician
In-place density	Roadway	10 tests per lot	Five companion tests per lot	Witness and test once per year by certified technician
Smoothness	Roadway	NA	NA	NA
Flat or elongated particles	Behind paver	One on first lot	One per project	Witness and test once per year by certified technician

NA = Not applicable
1 ton = 0.907 Mg

Louisiana

Quality Control

The contractor is responsible for developing the mix design for the project. The JMF must include the recommended formula, extracted gradation, and supporting design data. In addition, the contractor is responsible for providing a QC plan that must be approved by the engineer prior to construction and providing qualified personnel and equipment to conduct QC testing. The contractor must allow the engineer access to the laboratory to observe any and all testing procedures, calculations, test documentation, and plotting of results.

Mix design verification is evaluated on the following properties: gradation, asphalt content, percent crushed aggregate, MSG, corrected BSG @ N_{des}, percent air voids @ N_{des}, percent VMA @ N_{des}, percent VFA @ N_{des}, gyratory compaction curve @ N_{max}, slope of the compaction curve, and percent anti-strip.

Upon validation of the JMF, the validation parameters shall be used to control the process. The quality of the mixtures will be evaluated during two phases: (1) production of the mixture at the plant, and (2) hauling, laying, and compacting of the mixture. The quality of both phases must be evaluated continuously.

Acceptance

The Louisiana Department of Transportation and Development (DOTD) is responsible for all acceptance testing. Louisiana DOTD's acceptance tests are independent of the contractor's QC tests. The lot size is 5000 Mg, with 1000-Mg sublots unless otherwise specified by the engineer. Acceptance is based on the validation parameters used to verify the JMF. Acceptance testing for VMA, VFA, voids, maximum theoretical gravity, N_{ini}, N_{des}, N_{max}, percent anti-strip, quality of asphalt cement, extracted aggregate gradation, percent crushed aggregate, percent asphalt cement, and percent moisture in the mix will be conducted on the total lot quantity.

Pay Factors

Payment adjustments will be calculated based on specification limits and specified maximum variances from the JMF:

Plant Acceptance:

For: Percent voids
Percent VMA
Extracted aggregate gradation
#8 sieve
#200 sieve

PWL will be calculated for each of the parameters listed above. A mixture with any parameter below 40 PWL is subject to removal as directed by the engineer. The percent payment for each of the parameters will be calculated using the formula:

$$\text{Percent Payment} = 55 + 0.5PWL \tag{60}$$

For: Percent anti-strip additive

Payment adjustments for anti-strip additive will be based on the table below.

For: Asphalt cement

Payment adjustments for asphalt cement will be based on section 1012 of the specifications.

For: Percent payment for plant acceptance

All payment parameters will be averaged. The maximum payment factor shall be 100 percent.

Roadway Acceptance:
For: Pavement density

PWL will be calculated for the pavement density. A mixture with density results below 40 PWL is subject to removal as directed by the engineer. The percent payment for pavement density will be calculated using the formula:

$$\text{Percent Payment} = 55 + 0.5PWL \tag{61}$$

For: Surface tolerance (final wearing course travel lanes)

Payment adjustments for surface tolerance will be based on the table below.

For: Percent payment for roadway acceptance

All payment parameters will be averaged. The maximum payment factor shall be 100 percent.

Total Payment Factor:
The total payment factor for the mixture used for the project will be calculated using the following formula:

Total Percent Payment = (Percent payment for plant acceptance) + (Percent payment for roadway acceptance) – 100

Payment Adjustment Schedule: Smoothness and Percent Anti-Strip

Table 116. Louisiana payment adjustment schedule.

Surface Tolerance (inches/mi per lot)	Percentage of Contract Unit Price per Lot				
	100	98	95	80	50 or remove
Multi-lift new construction	0.0-3.0	—	3.1-4.0	4.1-6	More than 6.0
One- or two-lift overlays over planed surfaces	0.0-5.0	—	5.1-6.0	6.1-10	More than 10.0
Single-lift overlays over existing surface	0.0-8.0	—	8.1-10	10.1-15	More than 15.0
Anti-strip, percent below JMF	Within JMF	—	0.2 or less	More than 0.2	—

1 inch/mi = 15.8 mm/km

QC/QA Tests for Louisiana

Table 117. Louisiana QC/QA tests.

Test	Sampling Location	QC Testing Frequency by Contractor	Acceptance Testing Frequency by Louisiana DOTD
Gradation (extracted)	Mix from truck	Two per sublot	One per sublot
Percent crushed	Mix from truck	Two per sublot	One per sublot
Asphalt content	Mix from truck	Two per sublot	One per sublot
Percent VMA	Mix from truck	Two per sublot	Average of five per lot
Percent VFA	Mix from truck	Two per sublot	Average of five per lot
Air voids	Mix from truck	Two per sublot	Average of five per lot
Maximum theoretical gravity	Mix from truck	Two per sublot	One per lot
N_{ini}	Mix from truck	Two per sublot	One per lot
N_{des}	Mix from truck	Two per sublot	One per lot
N_{max}	Mix from truck	Two per sublot	One per lot
Percent anti-strip	Plant	Two per sublot	Two per lot
Asphalt properties	Plant	Two per sublot	One per lot
Percent moisture in mix	Mix from truck	Two per sublot	One per lot
Density	Roadway	Obtain samples for engineer	Five per sublot
Profile index	Roadway	One per day's placement	One per lot

Maine

Quality Control

The contractor is responsible for developing the mix design for the project, providing a QC plan that must be approved by the engineer prior to construction, and providing qualified personnel and equipment to conduct QC testing. The QC inspections and tests must be documented and provided to Maine DOT. The contractor must maintain records of all inspections and tests. These records shall indicate the nature, number, and type of deficiencies found; the quantities approved and rejected; and the nature of the corrective actions taken. The contractor must allow the engineer access to the laboratory to observe any and all testing procedures, calculations, test documentation, and plotting of results.

Mix design verification is evaluated on the following properties: gradation, asphalt content, theoretical MSG, CA angularity, FA angularity, air voids, VMA, VFA, dust-to-asphalt ratio, flat or elongated particles, percent moisture, and TSR.

The contractor must exercise control over the following properties: asphalt content, temperature of mix, temperature of mat, percent TMD, gradation, air voids, VMA, VFA, Rice specific gravity, CA angularity, and dust-to-asphalt ratio.

Quality Assurance

Maine DOT is responsible for QA and acceptance of the material. QA is accomplished using verification testing. Maine DOT's verification tests are independent of the contractor's QC tests. Lot size is usually 6000 Mg or less, with sublots of 1500 Mg or less.

Acceptance

The contractor's control tests are not used for acceptance. Acceptance is based on Maine DOT's verification tests. If the QC and QA test results do not agree, acceptance is based on referee testing for asphalt content, density, air voids, and VMA. The testing frequencies vary, depending on the property being measured.

Pay Factors

Pay factors are assigned for gradation, asphalt content, in-place density, air voids, and VMA. The determination of these factors uses quality-level analysis and price adjustment or *f* factors listed in the specification for the applicable property. Pay factors are assigned on a lot-by-lot basis. The composite pay factor is then computed by:

$$CPF = [f_1(PF_1) + f_2(PF_2) + \ldots. + f_j(PF_j)] / \sum f \tag{62}$$

QC/QA Tests for Maine

Table 118. Maine QC/QA tests.

Test	Sampling Location	Quality Control Testing Frequency by Contractor	Verification Testing Frequency by Maine DOT	Acceptance by Maine DOT
Gradation	Truck body or paver hopper	One per day minimum	One per 1500 Mg	Based on verification tests unless referee testing is required
Percent TMD	Mat behind all rollers	One per 125 Mg	One per 250 Mg	Based on verification tests unless referee testing is required
Dust-to-asphalt ratio	Truck body or paver hopper	One per 1500 Mg	One per 1500 Mg	Based on verification tests unless referee testing is required
Temperature of mix	Truck body or paver hopper	Six per day at plant and street	NA	NA
Temperature of mat	Roadway	Four per day	NA	NA
Asphalt content	Truck body or paver hopper	One per 750 Mg	One per 1500 Mg	Based on verification tests unless referee testing is required
Air voids	Truck body or paver hopper	One per 750 Mg	One per 1500 Mg	Based on verification tests unless referee testing is required
VMA	Truck body or paver hopper	One per 750 Mg	One per 1500 Mg	Based on verification tests unless referee testing is required
Rice specific gravity	Truck body or paver hopper	One per 750 Mg	NA	NA
CA angularity	Stockpile	One per 5000 Mg	NA	NA
Flat or elongated particles	Stockpile	One per 5000 Mg	NA	NA
Density (nuclear)	Roadway	One per 250 Mg	One per 250 Mg	Based on verification tests unless referee testing is required

NA = Not applicable

Minnesota

Quality Control

The contractor is responsible for developing the mix design in accordance with AASHTO TP-4, June 1997, or the Asphalt Institutes' *Superpave Mix Design Manual* (SP-2), such that it meets the requirements of the specification. In addition, the contractor is responsible for providing a QC plan that must be approved by the engineer prior to construction and providing qualified personnel and equipment to conduct QC testing. The contractor must allow the engineer access to the laboratory to observe any and all testing procedures, calculations, test documentation, and plotting of results.

Mix design verification is evaluated on the following properties: gradation; asphalt content; voids; VMA; TSR; compaction @ N_{des}, N_{ini}, and N_{max}; VFA; FA angularity; CA angularity; G_{mm}; G_{mb}; G_{sb}; mix moisture; and dust-to-asphalt ratio.

The contractor must exercise control over the following properties: asphalt content, gradation, voids, mix moisture, VMA, FA angularity, density, CA angularity, and percent G_{mm} @ N_{des}. In addition, the contractor must provide split samples to the engineer for verification testing. Testing frequency will be equal to the day's planned production divided by 1000, then rounded to the next whole number.

Quality Assurance

Minnesota DOT is responsible for QA and for the acceptance of the material. The assurance process is achieved through the use of verification tests conducted on split samples, observation of sampling and testing performed by the QC personnel, the taking of additional samples for testing, and the monitoring of the required QC summary sheets and control charts. Lot size varies based on the contract quantity. Verification tests are conducted once per day on all production parameters.

Acceptance

Acceptance is based on the QC tests as verified by Minnesota DOT's tests for gradation, asphalt content, maximum specific gravity, G_{mb}, air voids, VMA, and density.

Pay Factors

Pay factors are assigned for nonconforming mixes for the following properties: gradation, FA angularity, CA angularity, VMA, asphalt content, and G_{mm} @ N_{des}. For these characteristics, the lowest single payment applies according to the following table:

Payment Schedule

Table 119. Minnesota payment schedule.

Item	Percent Payment
FAA	95
CAA	95
Gradation (extracted)	90
VMA	85
Asphalt binder content	85
G_{mm} @ N_{des}	70

The pay factor for in-place density is determined on a lot-by-lot basis, with the number of lots determined by the following table:

Determination of Lots for Density

Table 120. Minnesota determination of lots for density.

Daily Production (tons)	Lots
0-300	1
301-600	2
601-1000	3
1001-1600	4
1601-3600	5
3601-5000	6
5000+	7

1 ton = 0.907 Mg

The pay factor is then computed using the following table:

Payment Schedule for Density

Table 121. Minnesota payment schedule for density.

Percentage of G_{mb} (< 100 mm from surface)	Percentage of G_{mb} (> 100 mm from surface)	Percent Payment
≥ 93.6	≥ 94.6	104
93.1-93.5	94.1-94.5	102
92.0-93.0	93.0-94.0	100
91.0-91.9	92.0-92.9	98
90.5-90.9	91.5-91.9	95
90.0-90.4	91.0-91.4	91
89.5-89.9	90.5-90.9	85
89.0-89.4	90.0-90.4	70
< 89.0	< 90.0	Remove and replace

The profile index is determined using a California-type profilograph. Profiles will be made for segments of 0.1 km. The following table determines the pay factor for smoothness:

Payment Schedule for Smoothness

Table 122. Minnesota payment schedule for smoothness.

mm/km per 0.1-km Segment	Dollars per Segment
0-13	110
14-25	90
26-38	65
39-79	0
80-92	-90
93-105	-180
106-118	-360
> 118	Corrective action

QC/QA Tests for Minnesota

Table 123. Minnesota QC/QA tests.

Test	Sampling Location	QC Testing Frequency by Contractor	Verification Test by Minnesota DOT	Acceptance by Minnesota DOT
Gradation (extracted)	Truck or paver	One day's planned production, divided by 1000, then rounded to the next whole number	One per day	Based on QC as verified by Minnesota DOT's tests
Asphalt content	Truck or paver	One day's planned production, divided by 1000, then rounded to the next whole number	One per day	Based on QC as verified by Minnesota DOT's tests
Air voids	Truck or paver	One day's planned production, divided by 1000, then rounded to the next whole number	One per day	Based on QC as verified by Minnesota DOT's tests
VMA	Truck or paver	One day's planned production, divided by 1000, then rounded to the next whole number	One per day	Based on QC as verified by Minnesota DOT's tests
G_{mm} @ N_{des}	Truck or paver	One day's planned production, divided by 1000, then rounded to the next whole number	One per day	Based on QC as verified by Minnesota DOT's tests
TSR	Plant	One per 10,000 tons	NA	NA
G_{sb}	Plant	One per 10,000 tons	NA	NA
CAA	Conveyor	One per day	One per day	Based on QC as verified by Minnesota DOT's tests
FAA	Conveyor	One per day	One per day	Based on QC as verified by Minnesota DOT's tests
Moisture content	Truck or paver	Engineer's discretion	Engineer's discretion	NA
Smoothness	Roadway	One pass per lane segment (0.1 km)	Observation of QC	Based on QC testing as verified by observation
Density (percentage of G_{mb})	Roadway	Three cores per lot	One companion core per lot	Based on QC as verified by Minnesota DOT's tests

NA = Not applicable
1 ton = 0.907 Mg

Mississippi

Quality Control

The contractor is responsible for developing the mix design for the project. In addition, the contractor is responsible for providing a QC plan that must be approved by the engineer prior to construction. The contractor must also provide qualified personnel and equipment to conduct QC testing. The contractor must allow the engineer access to the laboratory to observe any and all testing procedures, calculations, test documentation, and plotting of results.

Mix design verification is evaluated on the following properties: gradation, asphalt content, asphalt properties, voids @ N_{des}, VMA @ N_{des}, fractured faces, TSR, N_{des}, N_{ini}, N_{max}, stripping, RAP gradation, FA angularity, MSG, and dust-to-asphalt ratio.

The contractor must exercise control over the following properties: asphalt content, gradation, voids @ N_{des}, mix moisture, VMA @ N_{des}, FA angularity, density, fractured faces, MSG, stripping, and RAP gradation. In addition, the contractor must provide split samples to the engineer for verification testing. Sampling frequency for QC of voids, VMA, gradation, MSG, FA angularity, CA angularity, RAP gradation, stripping, and asphalt content is as follows:

Sampling Frequency

Table 124. Mississippi sampling frequency.

Total Estimated Production (tons)	Number of Tests
50-600	1
601-1500	2
1501-2700	3
2701+	4

1 ton = 0.907 Mg

Determination of the lots for density testing is as follows:

Lot Determination for Density

Table 125. Mississippi lot determination for density.

Daily Production (tons)	Number of Lots
0-300	1
301-600	2
601-1000	3
1001-1500	4
1501-2100	5
2101-2800	6
2800+	7

1 ton = 0.907 Mg

Quality Assurance

Mississippi DOT is responsible for QA and acceptance of the material. QA is accomplished by verification testing on split samples, testing additional samples at any time, and observing the QC tests. Lot size varies based on the contract quantity.

Acceptance

Acceptance is based on the QC tests as verified by Mississippi DOT's tests for total voids, VMA @ N_{des}, gradation, asphalt content, and density. Acceptance for smoothness is based on the observed contractor QC tests. Should the QC tests and the verification tests fail to meet the allowable differences, the engineer's test results will be used for acceptance.

Pay Factors

Pay factors are assigned for gradation, in-place density, total voids @ N_{des}, smoothness, and VMA @ N_{des}. Pay factors for these properties can be determined from the following tables:

Pay Factors for Mixture Quality

Table 126. Mississippi pay factors for mixture quality.*

Item	Produced in Warning Bands	Outside JMF limits (allowed to remain in place)
Gradation	0.90	0.75
Asphalt content	0.85	0.75
Total voids @ N_{des}	0.70	0.50
VMA @ N_{des}	0.90	0.75

* Minimum single pay factor applies.

Pay Factors for Density

Table 127. Mississippi pay factors for density.

Pay Factor	Lot Density (percentage of maximum density)
1.00	≥ 92.0
0.90	91.0-91.9
0.80	90.0-90.9

Pay Factors for Smoothness

Table 128. Mississippi pay factors for smoothness.

Profile Index (inch/mi per lot)	Pay Factor (percentage of unit bid price)
≤ 3	105
> 3-5	102
> 5-7	100
> 7-8	95
> 8-10	90
> 10	Unacceptable

1 inch/mi = 16 mm/km

QC/QA Tests for Mississippi

Table 129. Mississippi QC/QA tests.

Test	Sample Location	QC Testing Frequency by Contractor	Verification Testing Frequency by Mississippi DOT	Acceptance by Mississippi DOT
Binder sampling	Plant tank	One per 200,000 gal	≥ 10% of QC tests	NA
Binder content	Truck	See QC table 1	≥ 10% of QC tests	Based on verified QC by Mississippi DOT's tests
Mix gradation (extracted)	Truck	See QC table 1	≥ 10% of QC tests	Based on verified QC by Mississippi DOT's tests
MSG	Truck	See QC table 1	≥ 10% of QC tests	NA
Voids @ N_{des}	Truck	See QC table 1	≥ 10% of QC tests	Based on verified QC by Mississippi DOT's tests
Fractured faces	Stockpiles	One per day	≥ 10% of QC tests	NA
RAP gradation	Stockpiles	See QC table 1	≥ 10% of QC tests	NA
Moisture damage to mix	Plant mix	One per 2 weeks	≥ 10% of QC tests	NA
VMA @ N_{des}	Truck	See QC table 1	≥ 10% of QC tests	Based on verified QC by Mississippi DOT's tests
CA angularity	Stockpiles	See QC table 1	≥ 10% of QC tests	NA
FA angularity	Stockpiles	See QC table 1	≥ 10% of QC tests	NA
Density (nuclear)	Roadway	Two readings per lot	≥ 10% of QC tests	Based on verified QC by Mississippi DOT's tests
Smoothness	Roadway	One wheel path per lane per segment (0.1 mi)	Observation of QC tests	Verification by observation
Percent moisture in mix	Truck	Two per day	≥ 10% of QC tests	NA

NA = Not applicable
1 mi = 1.61 km, 1 gallon (gal) = 3.785 liters (L)

New York

Quality Control

The contractor is responsible for developing the mix design for the project. In addition, the contractor is responsible, through the mix manufacturer, for providing a QC plan that must be approved by the engineer prior to construction. The contractor must allow the engineer access to all testing procedures, calculations, test documentation, and plotting of results.

Mix design verification is evaluated on the following properties: gradation, asphalt content, compaction @ N_{ini}, N_{des}, N_{max}, VMA, VFA, dust-to-asphalt ratio, CA angularity, FA angularity, sand equivalent, and flat or elongated particles.

The contractor must exercise control over the following properties: asphalt content, gradation, MSG, BSG, air voids, VMA, VFA, moisture content, RAP extraction, mix temperature, and density. Additionally, the contractor must provide split samples for verification testing by the New York State DOT. The lot size is 1 day's production of a JMF with equal sublots not to exceed 1150 metric tons.

Quality Assurance

New York State DOT is responsible for QA and acceptance of the material. New York State DOT's assurance tests are independent of the contractor's QC tests. The contractor's control tests are used to assign a quality adjustment factor (QAF) for pay factors if the results agree with the verification test results. If New York State DOT's tests and the contractor's tests are not within specified tolerances, referee testing is conducted on independent samples and these results will be used for the determination of the QAF.

Acceptance

Acceptance for density is based on the PWL of the percent MTD. Acceptance for gradation, air voids, mix moisture, asphalt content, VMA, VFA, MSG, and BSG is based on the contractor's QC tests if the results are within specified tolerances of New York State DOT's tests.

Pay Factors

Pay factors are assigned for gradation, asphalt content, in-place density, and air voids. The method for determining the pay factors is currently being researched.

QC/QA Tests for New York

Table 130. New York QC/QA tests.

Test	QC Testing by Contractor	Assurance Testing by New York State DOT	Acceptance by New York State DOT
Gradation (extracted)	One per every other sublot, minimum of one per day	One per lot minimum	Based on verification of QC by New York State DOT's tests
Aggregate moisture	One per every other sublot, minimum of two per day	One per lot minimum	NA
Air voids	One per sublot	One per lot minimum	Based on verification of QC by New York State DOT's tests
Wet analysis of percent passing #200	One per week	NA	NA
Mix moisture	As required by New York State DOT	One per lot minimum	Based on verification of QC by New York State DOT's tests
Asphalt content	Four times per day minimum	One per lot minimum	Based on verification of QC by New York State DOT's tests
RAP moisture	Two per week	NA	NA
RAP extraction	Two per week	NA	NA
Binder test	Two per day	One per lot minimum	Based on verification of QC by New York State DOT's tests
VMA	One per sublot	One per lot minimum	Based on verification of QC by New York State DOT's tests
VFA	One per sublot	One per lot minimum	Based on verification of QC by New York State DOT's tests
Density	Four cores per 1800 lane-meters	Average of four per lot	Based on verification of QC by New York State DOT's tests
MSG	One per day	One per lot minimum	Based on verification of QC by New York State DOT's tests
BSG	One per day	One per lot minimum	Based on verification of QC by New York State DOT's tests
Mix temperature	Four times per day minimum	NA	NA

NA = Not applicable

North Carolina

Quality Control

The contractor is responsible for developing the mix design for the project and providing a QC plan that must be approved by the engineer prior to construction. The QC program will include process control inspection, sampling and testing, and necessary adjustments in the process that are related to the production of a pavement that meets all of the requirements of the specification. The contractor must also provide qualified personnel and equipment to conduct QC testing. The contractor must allow the engineer access to the laboratory to observe any and all testing procedures, calculations, test documentation, and plotting of results.

The contractor must exercise control over the following properties: asphalt content, gradation, air voids @ N_{des}, VMA @ N_{des}, VFA @ N_{des}, G_{mm}, G_{mb} @ N_{des}, dust-to-asphalt ratio, percentage of MSG @ N_{ini}, and aggregate moisture content. In addition, the contractor must provide split samples to the engineer for assurance testing. Testing frequency is as follows:

Testing Frequency

Table 131. North Carolina testing frequency.

Daily Production (metric tons)	Number of Tests
80-1000	One
1001-2500	Two
2500+ (in 1500 increments)	One per 1500 metric tons

Quality Assurance

North Carolina DOT is responsible for QA and for acceptance of the material. The verification tests are conducted on split samples. Testing frequency is equal to or greater than 10 percent of the contractor's control testing. QA is accomplished by conducting verification tests on split samples, observing tests conducted by the contractor, monitoring control charts, and conducting assurance testing on samples independent of the QC tests.

Acceptance

Acceptance is based on the QC tests as verified by North Carolina DOT's tests for asphalt content, G_{mm}, G_{mb} @ N_{des}, and density.

Pay Factors

Pay factors are assigned for gradation, binder content, air voids, VMA @ N_{des}, density, and profile index. Pay factors are determined as follows:

Payment for Mix Produced in Warning Bands*

Table 132. North Carolina payment for mix produced in warning bands.

Item	Percentage of Bid Price for Mix
Gradation	90
Binder content	85
Air voids	70
VMA @ N_{des}	90

* Minimum single payment applies.

The pay factor for density is determined using the following formula:

$$PF = 100 - 10(D)^{1.465} \qquad (63)$$

where: PF = pay factor computed to the nearest 0.1 percent
D = deficiency of the lot average density, not to exceed 3.0 percent

The pay factor for smoothness is determined on a lot-by-lot basis, where a lot is 25 test sections of 100 ft (30 m) of single lane. Measurements are made using a North Carolina Hearne Straightedge. The pay factor is assigned as follows:

Table 133. North Carolina pay factor.

Smoothness Index	Acceptance Category	Corrective Action	Pay Adjustments
0.0-9.9	Acceptable	None	+$300
10-20	Acceptable	None	+$100
30-40	Acceptable	None	No adjustment
50-60	Acceptable	Allowed	-$300
Any other number	Unacceptable	Required	-$600

QC/QA Tests for North Carolina

Table 134. North Carolina QC/QA tests.

Test	Sampling Location	QC Frequency by Contractor	Verification Frequency by North Carolina DOT	Acceptance by North Carolina DOT
Gradation (extracted)	Truck at plant	See QC table 1	≥ 10% of QC testing	Based on QC as verified by North Carolina DOT's tests
Asphalt content	Truck at plant	See QC table 1	≥ 10% of QC testing	Based on QC as verified by North Carolina DOT's tests
Air voids	Truck at plant	See QC table 1	≥ 10% of QC testing	Based on QC as verified by North Carolina DOT's tests
VMA	Truck at plant	See QC table 1	≥ 10% of QC testing	Based on QC as verified by North Carolina DOT's tests
G_{mm}	Truck at plant	See QC table 1	NA	NA
G_{mb}	Truck at plant	See QC table 1	NA	NA
VFA	Truck at plant	See QC table 1	NA	NA
Dust-to-asphalt ratio	Truck at plant	See QC table 1	NA	NA
Percent MSG	Truck at plant	See QC table 1	NA	NA
Aggregate moisture content	Cold feed	One per day	NA	NA
Smoothness	Roadway	Test each 100 ft of single lane	Observe QC test	As verified by observation
Density	Roadway	One test per 1500 linear-feet	≥ 10% of QC testing	Based on QC as verified by North Carolina DOT's tests

NA = Not applicable
1 ft = 0.305 m

APPENDIX D: SUMMARY OF AGENCY PCC SPECIFICATIONS RECEIVED

Kansas

Quality Control

The contractor is responsible for providing a QC plan that must be approved by the engineer prior to construction and providing qualified personnel and equipment to conduct QC testing. The contractor must allow the engineer access to the laboratory to observe any and all testing procedures, calculations, test documentation, and plotting of results. At a minimum, values for percent air, slump, unit weight, and gradation must be plotted on control charts. QC tests include aggregate gradation, slump, air content, unit weight/yield, compressive strength, flexural strength, material passing #200, percent moisture in aggregate, temperature, and density of fresh concrete. Cores are to be tested once per each sublot, where a sublot is the day's lot production divided into five sublots of equal surface area.

Quality Assurance

QA is the responsibility of the engineer and is accomplished through verification testing by Kansas DOT. The engineer conducts verification testing on aggregate gradation, slump, air content, unit weight/yield, compressive strength, flexural strength, material passing #200, percent moisture in aggregate, temperature, and density of fresh concrete.

Acceptance

Acceptance or rejection of material is based on the contractor's QC tests as verified by the engineer's tests. Should the results of the two tests disagree, the engineer's results will be used to determine acceptance or rejection. If the test results are disputed, referee testing will be conducted on additional samples by an independent laboratory. The referee test results will then be used for acceptance.

Pay Factors

Pay factors are assigned for smoothness, compressive strength, and thickness. The smoothness pay factor is based on the California profilograph. Values for this pay factor were not provided. Pay factors for thickness and compressive strength are based on a quality index for each characteristic. The quality index (Q_i) is calculated as follows:

$$Q_i = (X_{avg} - LSL)/S \tag{64}$$

where:
- X_{avg} = average measured thickness or compressive strength of all cores in the lot
- LSL = lower specification limit
- S = standard deviation of all thickness or compressive strength samples for the lot

The quality index is then used to calculate a pay adjustment factor according to the following table:

Concrete Thickness and Compressive Strength Pay Adjustment

Table 135. Kansas concrete thickness and comprehensive strength pay adjustment.

Quality Index	Pay Adjustment Factor
> 1.71	103%
1.56-1.71	102%
1.41-1.55	101%
1.26-1.40	100%
0.00-1.25	*
< 0	50%

*Calculated as $P = 60 + (Q \times 32)$

A composite pay factor for the lot is then calculated as $P_c = (P_t \times P_s)/100$.

QC/QA Tests for Kansas

Table 136. Kansas QC/QA tests.

Test	Sampling Location	QC Testing by Contractor	Verification Testing by Kansas DOT
Aggregate gradation	Feed bins	One test per 1000 tons	One per 30,000 tons
Slump	Truck	One per 500 yd^3 or minimum of one per day	One per day
Air content (plastic concrete)	Truck	One per 500 yd^3 or minimum of one per day	One per day
Cores	Roadway	One per sublot	One per lot
Temperature	Truck	One per 500 yd^3	One per day
Beams	Truck	As required for opening to traffic	One set per week
Unit weight	Truck	One per 500 yd^3 or minimum of one per day	One per day
Material passing #200 (washed)	Feed bins	One test per 1000 tons	One per week
Smoothness	Roadway	NA	NA
Thickness	Roadway	One per sublot	One per lot
Density of fresh concrete	Roadway	One complete transverse profile initially, then one per ½ day	One per week

NA = Not applicable
1 cubic yard (yd^3) = $0.765 m^3$, 1 ton = 0.907 Mg

Illinois

Quality Control
The contractor is responsible for providing a QC plan that must be approved by the engineer prior to construction and providing qualified personnel and equipment to conduct QC testing. The contractor must allow the engineer access to the laboratory to observe any and all testing procedures, calculations, test documentation, and plotting of results. At a minimum, the contractor must have a level 1 PCC technician at the job site, as well as a level 2 PCC technician at the plant during mixture production and placement. QC tests include aggregate gradation, slump, air content, unit weight/yield, temperature, and compressive or flexural strength.

Quality Assurance
The engineer is responsible for all QA testing on split samples. The split sample is one of two equal portions of a field sample. At the engineer's discretion, he/she may also obtain independent field samples for testing. The results of the QA tests will be made available to the contractor as soon as they are completed. In the event that the contractor's or the engineer's tests for air content and strength are not within specification limits, referee testing will be conducted by Illinois DOT or by a mutually agreed upon laboratory at the contractor's option.

Acceptance
Acceptance of the material is based on: contractor's compliance with all contract documents for QC, validation of the contractor's QC test results by comparison with the engineer's QA tests, comparison of the engineer's QA test results with the specification limits using samples independently obtained by the engineer, and referee test results for unresolved failing tests obtained by the contractor or the engineer.

Pay Factors
Pay factors are assigned for smoothness and thickness.

QC/QA Tests for Illinois

Table 137. Illinois QC/QA tests.

Test	Sampling Location	QC Testing Frequency by Contractor	Verification Testing Frequency by Illinois DOT (split sample)
Aggregate gradation	At plant from bins or stockpiles	One per 2500 yd^3	First test performed by the contractor at the beginning of the project, and then a minimum of 10% of the tests required by the contractor
Slump	Job site	One per 500 yd^3 or minimum of one per day	First three tests performed by the contractor at the beginning of the project, and then a minimum of 10% of the tests required by the contractor
Air content (plastic concrete)	Job site	One per 100 yd^3 or minimum of one per day	First three tests performed by the contractor at the beginning of the project, and then a minimum of 10% of the tests required by the contractor
Strength	Laboratory	One per 1250 yd^3 or minimum of one per day	First test performed by the contractor at the beginning of the project, and then a minimum of 10% of the tests required by the contractor
Smoothness	Job site	NA	NA
Thickness	Job site	NA	NA
Temperature	Plant	As needed to control production	NA

NA = Not applicable
1 yd^3 = 0.765 m^3

Iowa

Quality Control

The contractor is responsible for providing a QC plan that must be approved by the engineer prior to construction and providing qualified personnel and equipment to conduct QC testing. The contractor must allow the engineer access to the laboratory to observe any and all testing procedures, calculations, test documentation, and plotting of results. At a minimum, the contractor must perform QC tests on the unit weight of plastic concrete, aggregate gradation, flexural strength, air content, slump, and water-to-cement ratio.

Quality Assurance

The engineer is responsible for QA and acceptance of the material. QA is accomplished by conducting verification sampling and testing, observing the QC sampling and testing, taking additional samples for testing at any time, and monitoring control charts. Results of the verification tests will be made available to the contractor as soon as they are completed. Verification testing frequency will be at the discretion of the engineer, but must not be less than 10 percent of the QC tests. Verification testing will be conducted on those properties tested for QC.

Acceptance

Acceptance of the material will be based on the contractor's QC test results as long as the results agree with Iowa DOT's test results. Acceptance for thickness will be based on samples tested by the engineer. Coring for thickness will be on a lot-by-lot basis, where 1 day's production is considered a lot, which is divided into three equal sublots for testing.

Pay Factors

Pay factors are assigned for smoothness, thickness, and flexural strength. The pay factor for smoothness was not provided. The pay factor for thickness is determined using the following table:

Table 138. Iowa pay factors for thickness.

Thickness Index* Range (mm)	Percent Payment	Thickness Index* Range (mm)	Percent Payment
0.00 or more	103	–13.98 to –15.24	91
–0.01 to –1.27	102	–15.25 to –16.51	90
–1.28 to –2.54	101	–16.52 to –17.78	89
–2.55 to –3.81	100	–17.79 to –19.05	88
–3.82 to –5.08	99	–19.06 to –20.32	87
–5.09 to –6.35	98	–20.33 to –21.59	86
–6.36 to –7.62	97	–21.60 to –22.86	85
–7.63 to –8.89	96	–22.87 to –24.13	84
–8.90 to –10.16	95	–24.14 to –25.40	83
–10.17 to –11.43	94	–25.41 to –26.67	82
–11.44 to –12.70	93	–26.68 to –27.94	81
–12.71 to –13.97	92	–27.95 or less	80

* The thickness index is determined as follows: $TI = (X_{avg} - S) - T$, where X_{avg} = mean core length for the section, S = core length standard deviation for the section, and T = design thickness. Pavement, represented by cores, which is deficient from the design thickness by 25 mm or greater shall be replaced. Payment for thickness will be based on the percentage of the contract price.

The pay factor for flexural strength is determined by subtracting 1 standard deviation from the mean strength according to the following table. The average strength is based on three beams.

Table 139. Iowa pay factors for flexural strength.

Strength (lbf/inch^2)	Pay Factor
< 450	Remove
450-474	70
475-499	75
500-524	80
525-549	85
550-559	90
560-569	92
570-579	94
580-589	96
590-599	98
600-624	100
625-649	101
650-674	102
≥ 675	103

1 lbf/inch^2 = 6.89 kPa

The composite pay factor was not provided; however, it appears to be the individual pay factors multiplied together times the contract price.

QC/QA Tests for Iowa

Table 140. Iowa QC/QA tests.

Test	Sampling Location	QC Testing Frequency by Contractor	Verification Testing Frequency by Iowa DOT	Acceptance by Iowa DOT
Aggregate gradation	Feed bins	Two per day	At least 10% of QC tests	Verification of QC by Iowa DOT's tests
Air content (plastic concrete)	Truck	First load and then one per 500 yd^3	At least 10% of QC tests	Verification of QC by Iowa DOT's tests
Water-to-cement ratio	Truck	Two per day	At least 10% of QC tests	Verification of QC by Iowa DOT's tests
Unit weight of plastic concrete	Truck	One per day	At least 10% of QC tests	Verification of QC by Iowa DOT's tests
Thickness	Roadway	NA	NA	Five cores per sublot
28-day flexural strength	Truck	One per 1000 yd^3	At least 10% of QC tests	One set of three beams as determined by engineer

NA = Not applicable
1 yd^3 = 0.765 m^3

New Jersey

Quality Control

All plants producing concrete for New Jersey DOT projects shall have a QC plan in place, as outlined in "Requirements for a Portland Cement Concrete Quality Control Plan," June 25, 1998. A QC technician who is certified by the American Concrete Institute as grade I must be available during the production of concrete for any New Jersey DOT project. QC tests for slump, air content, compressive strength, aggregate gradation, and temperature are conducted at the plant. Concrete provided by any producer who has met applicable design, control, and acceptance testing requirements will be presumed to be in compliance with New Jersey DOT's standards at the time of delivery. However, the presumption will not waive the engineer's right to impose pay adjustments.

Acceptance

Acceptance of the material is determined through acceptance testing by New Jersey DOT. Acceptance will be based on the results of the slump, strength, and air content tests. Acceptance for compressive strength, slump, and air content will be on a lot-by-lot basis, with a lot being equal to 1 day's production. Acceptance tests will be conducted at a frequency of five per lot.

Pay Factors

Pay factors are assigned for air content and compressive strength. The pay factor for compressive strength and air content will be determined by the percent defective within the lot. Percent pay adjustments (PPA) are calculated using the following two equations:

$$\text{For } PD < 50: \quad PPA = 3.0 - 0.3PD \tag{65}$$

$$\text{For } PD \geq 50: \quad PPA = 26.0 - 0.76PD \tag{66}$$

North Carolina

Quality Control

The contractor is responsible for providing a QC plan that must be approved by the engineer prior to construction. During the preconstruction meeting, the contractor and the engineer will determine the types and frequencies of testing that are required to produce materials that meet the specifications. The contractor must also provide qualified personnel and equipment to conduct QC testing. The contractor must allow the engineer access to the laboratory to observe any and all testing procedures, calculations, test documentation, and plotting of results. At a minimum, the contractor must have a technician who is certified to develop the concrete design, control the gradation and quality of the aggregates, and perform the required tests. QC testing must be conducted for aggregate gradation, air content, slump, compressive strength, profile index, and thickness.

Acceptance

Acceptance of the material is determined through acceptance testing by North Carolina DOT. Acceptance will be based on the results of the slump, flexural strength, thickness, profile index, and air content tests. Acceptance for slump and air content will be determined at the point of placement at a frequency determined by the engineer. A sample will be obtained immediately after the concrete has been discharged onto the road. Acceptance for thickness and flexural strength will be determined on a lot-by-lot basis, where a lot is 5333.3 yd^3 and subsequently divided into four equal sublots. For each sublot, a set of two beams shall be used to test flexural strength. Additionally, two cores shall be taken from each lot to test for thickness. Acceptance testing of the longitudinal profile of the finished pavement must be performed by the contractor in the presence of the engineer.

Pay Factors

Pay factors are assigned for profile index, thickness, and flexural strength. The pay factor for the profile index was not provided. The pay factor for thickness is determined from the following table:

Table 141. North Carolina pay factors for thickness.

Deficiency (inches)	Percentage of Price Allowed
0.00-0.20	100
0.21-0.30	80
0.31-0.40	72
0.41-0.50	68
0.51-0.75	57
0.76-1.00	50

1 inch = 25.4 mm

The pay factor for flexural strength is determined from the following table:

Table 142. North Carolina pay factors for flexural strength.

Pay Factor Level	Lower Acceptance Limits			
	Number of Tests in Lot			
	3	4	5	6
1.05	550 + 0.60*R*	550 + 0.66*R*	550 + 0.66*R*	550 + 0.66*R*
1.00	550 + 0.58*R*	550 + 0.54*R*	550 + 0.50*R*	550 + 0.46*R*
0.95	550 + 0.36*R*	550 + 0.27*R*	550 + 0.23*R*	550 + 0.21*R*
0.70	550	550	550	550

Note: *R* is the difference between the high and low tests in the lot. If multiple deficiencies occur, the payment is determined by successively multiplying the contract price by the appropriate factor indicated for each deficiency.

QC/QA Tests for North Carolina

Table 143. North Carolina QC/QA tests.

Test	Sampling Location	QC Testing Frequency by Contractor	Acceptance Testing Frequency by North Carolina DOT
Aggregate gradation	Feed bins	Contractor's discretion as needed for control	NA
Slump	Roadway after discharge	Two per day	Engineer's discretion
Air content (plastic)	Roadway after discharge	Two per day	Engineer's discretion
Profile index	Roadway	Three profiles: Two at 3.5 ft inside the outer wheel path and one along the longitudinal joint	Based on QC test
Thickness	Roadway (coring)	Contractor's discretion as needed for control	Two cores per lot
Flexural strength	Truck	NA	Two beams per sublot (average)
Compressive strength	Roadway (coring)	Contractor's discretion as needed for control	NA

NA = Not applicable
1 ft = 0.305 m

Oregon

Quality Control

The contractor is responsible for providing a QC plan that must be approved by the engineer prior to construction and providing qualified personnel and equipment to conduct QC testing. The contractor must allow the engineer access to the laboratory to observe any and all testing procedures, calculations, test documentation, and plotting of results. At a minimum, the contractor must have a concrete control technician (CCT) at the plant or job site and a QC technician (QCT) at the job site. QC tests include aggregate gradation, slump, air content, water-to-cement ratio, unit weight/yield, temperature, and compressive strength. QC tests will be conducted on a lot-by-lot basis, where a lot is the total amount of concrete produced for each mix design. This lot will then be divided and tested by the sublot, where a sublot is equal to 75 m^3 of material.

Quality Assurance

The engineer is responsible for all QA. Assurance is accomplished through verification testing, observation of contractor sampling and testing, and additional sampling and testing conducted by the engineer. The results of the verification tests will be made available to the contractor as soon as they are completed. In the event that the contractor's or the engineer's tests are not within specification limits, they shall immediately work together to resolve the difference to avoid having the material rejected as not meeting specifications. Verification testing will be conducted on a lot-by-lot basis, where a lot is the total amount of concrete produced for each mix design.

Acceptance

Acceptance of the material is based on verification of the QC tests by Oregon DOT's tests. For aggregate acceptance, verification of the contractor's test values through check tests completed by the engineer will be used. For plastic concrete, acceptance is based on the QC tests performed by the contractor and on tests conducted by the engineer. The engineer may observe the QC tests performed by the contractor and/or conduct tests on all plastic concrete. Hardened concrete will be accepted based on the statistical analysis of the 28-day strength tests of cylinders cast by the engineer.

Pay Factors

Pay factors are assigned for smoothness, thickness, and compressive strength. The pay factor for smoothness is determined by the following equation:

$$\text{Bonus} = 0.00038(80 - PI)(\text{Quantity})(\text{Unit price}) \qquad (67)$$

where:
- PI = average of the two wheel path profiles for the segment
- Quantity = area in m^2 represented by the segment
- Unit price = price for the concrete shown in the bid schedule

The pay factor for thickness is determined using the following table. No additional payment will be made for pavement exceeding the minimum specification.

Table 144. Oregon pay factors for thickness.

Deficiency in Thickness (mm)	Proportional Part of Contract Unit Price Allowed
0.0-5.0	100%
5.1-7.6	83%
7.7-10.1	76%
10.2-12.7	73%
12.8-19.0	63%
19.1-25.0	59%

The pay factor for compressive strength is based on statistical analysis to determine the total percent within the specification limits. This is then used to determine the pay factor for strength. In no case shall payment exceed the contract price. For pay factors less than 1.00, a pay reduction is calculated as follows:

$$\text{Payment} = 0.3(PF - 1.0)(\text{Contract bid price}) \tag{68}$$

A weighted pay factor (*WPF*) is determined by multiplying *PF* by a weighting factor f_I provided in the contract. The method for determining a composite pay factor (CPF) for these items uses the following formula:

$$CPF = \frac{\sum WPF}{\sum f_I} \tag{69}$$

where: *WPF* = weighted pay factor
f_I = weighting factor

QC/QA Tests for Oregon

Table 145. Oregon QA/QC tests.

Test	Sample Location	QC Testing Frequency by Contractor	Verification Testing by Oregon DOT	Acceptance by Oregon DOT
Aggregate gradation	Feed bins	One per shift or 500 Mg	One per shift or 500 Mg	Verification of QC by Oregon DOT's test
Slump	Roadway after discharge	One per sublot	One per lot	Verification of QC by Oregon DOT's test
Air content (plastic)	Roadway after discharge	One per sublot	One per lot	Verification of QC by Oregon DOT's test
Water-to-cement ratio	Roadway after discharge	One per sublot	One per lot	Verification of QC by Oregon DOT's test
Yield	Roadway after discharge	One per sublot	One per lot	Verification of QC by Oregon DOT's test
Thickness	Roadway	One sticking of plastic concrete per 60 lane-meters	Observation of QC test	Observation of QC test
Compressive strength	Cast in the field	NA	NA	One set of four cylinders per sublot
Smoothness	Roadway	One pass per wheel path per 200-m segment	Observation of QC test	Observation of QC test
Temperature	Roadway after discharge	One per sublot	One per lot	NA

NA = Not applicable

Pennsylvania

Quality Control
The contractor is responsible for providing a QC plan that must be approved by the engineer prior to construction and providing qualified personnel and equipment to conduct QC testing. The contractor must allow the engineer access to the laboratory to observe any and all testing procedures, calculations, test documentation, and plotting of results. At a minimum, the contractor must have a certified technician to develop the concrete design, control the gradation and quality of the aggregates, and perform the required tests. QC testing must be conducted for aggregate gradation, air content, slump, compressive strength, profile index, and thickness.

Quality Assurance
QA is accomplished through independent sampling and testing by Pennsylvania DOT for air content, strength, and pavement thickness. Independent assurance testing is conducted to review the QC testing and to check the accuracy of the equipment used for acceptance testing.

Acceptance
Acceptance of the material is determined through acceptance testing by Pennsylvania DOT. Acceptance will be based on the results of the slump, strength, thickness, and air content tests. Acceptance for compressive strength and air content will be on a lot-by-lot basis, with testing conducted on sublots of 1400 yd^2 (1170 m^2). Acceptance for smoothness will be based on the QC profilograph test.

Pay Factors
Pay factors are assigned for profilograph, thickness, air content, and compressive strength tests. The pay factor for compressive strength and air content will be determined by the PWL for the lot. The pay factor for thickness will be determined by coring. The profile index pay factor is determined using a profilograph and assessing the index in inches per mile per lot. After determining the individual pay factors for each of the above characteristics, the lot payment is calculated using the following equation:

$$L_p = C_p\big([(2P_s + 2P_d + P_a)/500] + [(P_p - 100)/100]\big) \qquad (70)$$

where:
- L_p = lot payment
- C_p = contract price per lot
- P_s = payment percentage of the contract price for strength
- P_d = payment percentage of the contract price for depth
- P_a = payment percentage of the contract price for air content
- P_p = payment percentage of the contract price for the profile

QC/QA Tests for Pennsylvania

Table 146. Pennsylvania QA/QC tests.

Test	Sample Location	QC Testing	Assurance Testing	Acceptance Testing
Aggregate gradation	Feed bins	Varies: Established in QC plan	NA	NA
Slump	Roadway after discharge	One per 200 yd^3	NA	One per sublot
Air content (plastic)	Roadway after discharge	One per 200 yd^3	One per two lots	One per sublot
Profilograph	Roadway	Test as soon as concrete is cured	NA	Based on QC test
Thickness	Roadway	One core per sublot	One core per two lots	Four cores per lot
Compressive strength	Cast in the field	Two cylinders per sublot	One cylinder per two lots	Four cylinders per lot

NA = Not applicable
1 yd^3 = 0.765 m^3

Texas

Quality Control

The contractor is responsible for providing a QC plan that must be approved by the engineer prior to construction and providing qualified personnel and equipment to conduct QC testing. The contractor must allow the engineer access to the laboratory to observe any and all testing procedures, calculations, test documentation, and plotting of results. To complete the required testing, the contractor must provide technicians certified at levels I through IV, depending on the test being conducted. The standard lot size is 6000 yd^2 (5000 m^2) of surface area of PCC pavement of the same thickness. A sublot is equal to one-fifth of the surface area of a lot. If the final lot is 3000yd^2 (2500 m^2) or greater, it will be considered a full lot. QC must be exercised on all properties contained in the table entitled QC/QA Tests for Texas.

Quality Assurance

QA is accomplished through verification testing conducted by Texas DOT. Independent assurance testing is conducted to review the QC testing and to check the accuracy of the equipment used for verification testing.

Acceptance

Acceptance of the material is based on the QC tests and the verification tests. If the contractor's QC test results and the engineer's verification test results are not within tolerances, referee testing will be conducted. If referee tests are requested, the referee test results shall govern.

Pay Factors

Pay factors are assigned for ride quality, thickness, and flexural strength. The pay factor for flexural strength is based on the following table:

Table 147. Texas pay factors for flexural strength.

Average Flexural Strength (Three Beams) (lbf/inch^2)	Strength Adjustment Factor (SF)
≥ 555	1.0000
550	0.9695
545	0.9397
540	0.9106
535	0.8820
530	0.8542
525	0.8269
< 525	Remove or remain with no payment

1 lbf/inch^2 = 6.89 kPa

The pay factor for ride quality was not provided. The pay factor for thickness is determined using the following table:

Table 148. Texas pay factors for thickness.

Thickness Deficiency (inches)	Thickness Adjustment Factor (TF)
> 0.00-0.20	1.00
> 0.20-0.30	0.80
> 0.30-0.40	0.72
> 0.40-0.50	0.68
> 0.50-0.75	0.57

The pay adjustment factor for each lot is calculated as $AF = (SF)(TF)$. The pay adjustment for each lot is then calculated as:

$$\text{Payment} = (BP)(AF)(Q) \tag{71}$$

where: BP = bid price
Q = lot quantity of acceptable pavement

QC/QA Tests for Texas

Table 149. Texas QC/QA tests.

Test	Sample Location	QC Testing by Contractor	Verification Testing by Texas DOT	Acceptance by Texas DOT
Mix temperature	Plant	One per sublot	One per day	Based on QC, verification, and referee, if required
Air content (plastic)	Roadway after discharge	One per sublot	1 per 10 QC tests	Based on QC, verification, and referee, if required
Concrete unit weight	Roadway after discharge	One per sublot	1 per 10 QC tests	Based on QC, verification, and referee, if required
Making/curing strength specimens	Roadway after discharge	One per sublot	1 set per 10 sublots	NA
7-day flexural strength	Roadway after discharge	NA	One per sublot	Based on verification
Coarse aggregate gradation	Stockpile	One per stockpile per day	1 per 10 QC tests	Based on QC, verification, and referee, if required
Coarse aggregate loss by decantation	Stockpile	One per 5 PCC production days	1 per 10 QC tests	Based on QC, verification, and referee, if required
Fine aggregate gradation	Stockpile	One per stockpile per day	1 per 10 QC tests	Based on QC, verification, and referee, if required
Fineness modulus	Stockpile	One per stockpile per day	1 per 10 QC tests	Based on QC, verification, and referee, if required
Fine aggregate organic impurity	Stockpile	One per 5 PCC production days	1 per 10 QC tests	Based on QC, verification, and referee, if required
Sand equivalent	Stockpile	One per 5 PCC production days	1 per 10 QC tests	Based on QC, verification, and referee, if required
Aggregate moisture content	Stockpile	One per stockpile per day	1 per 10 QC tests	Based on QC, verification, and referee, if required
Water-to-cement ratio	Roadway after discharge	One per batch	Review plant printout and calibration	Based on QC, verification, and referee, if required
Cement factor	Roadway after discharge	One per batch	Review plant printout and calibration	Based on QC, verification, and referee, if required
Admixture dosage	Roadway after discharge	One per batch	Review plant printout and calibration	Based on QC, verification, and referee, if required
Coring for thickness	Roadway	One per sublot	NA	NA
Measuring pavement thickness	Roadway (plastic)	NA	One per sublot	Based on QC, verification, and referee, if required

NA = Not applicable

Wisconsin

Quality Control

The contractor is responsible for providing a QC plan that must be approved by the engineer prior to construction and providing qualified personnel and equipment to conduct QC testing. The contractor must allow the engineer access to the laboratory to observe any and all testing procedures, calculations, test documentation, and plotting of results. At a minimum, the contractor must have the following certified personnel: aggregate technician 1, PCC technician 1A, PCC technician 2, concrete compressive strength tester, and profilograph operator 1. QC testing must be conducted on the following: aggregate gradation, material passing #200, aggregate moisture, air content, slump, temperature, thickness, profilograph, and compressive strength.

Quality Assurance

QA is accomplished through verification testing by Wisconsin DOT for air content, strength, and pavement thickness. Independent assurance testing is conducted to review the QC testing and the verification testing. The independent assurance review will be done in accordance with Wisconsin DOT's Independent Assurance program and may include: split-sample testing, proficiency sample testing, witnessing of sampling and testing, review of control charts, and requesting additional samples for testing.

Acceptance

Acceptance of the material is based on the contractor's QC test results. This method of acceptance and payment will continue until it can be shown through verification or the dispute resolution process that the contractor's test results are in error. Acceptance will be based on the results of the strength, thickness, and air content tests.

Pay Factors

Pay factors are assigned for profilograph, thickness, and compressive strength. The pay factor for compressive strength will be determined by the lot and will be based on the contractor's QC cylinders fabricated for each sublot. A lot is considered to be 1 day's production divided into equal sublots not to exceed 500 yd^3 (380 m^3). The pay adjustment will be based on the lot's average strength minus 1 standard deviation. The factor is a dollars per yd^2 adjustment. The factor is determined from the following table:

Table 150. Wisconsin pay factors for compressive strength.

Strength ($lbf/inch^2$)	Pay Adjustment ($\$/yd^2$)
≤ 2850	–0.552
2850 to 2950	–0.527
2950 to 3050	–0.452
3050 to 3150	–0.385
3150 to 3250	–0.309
3250 to 3350	–0.234
3350 to 3450	–0.167
3450 to 3550	–0.109
3550 to 3650	–0.050
3650 to 3750	–0.000
3750 to 3850	0.067
3850 to 3950	0.125
3950 to 4050	0.167
4050 to 4150	0.201
4150 to 4250	0.226
4250 to 4350	0.242
4350 to 4450	0.259
4450 to 4550	0.268
4550 to 4650	0.268
≥ 4650	0.276

1 $lbf/inch^2$ = 6.89 kPa, 1 yd^2 = 0.836 m^2

The pay factor for the profile index will be determined according to the following table:

Table 151. Wisconsin pay factors for the profile index.

Profile Index (inches/mi)	Pay Adjustment per 0.1-mi Section per Lane
< 19.0	+$585
19.0-25.2	+$350
25.3-44.3	$0
44.4-50.6	–$230
≥ 50.7	–$940

1 mi = 1.61 km, 1 inch/mi = 15.8 mm/km

The pay factor for thickness will be determined according to the following table:

Table 152. Wisconsin pay factors for thickness.

Average Thickness Deficiency (inches)	Pay Adjustment per 250-ft Lane Length Unit
0-⅜	$0
> ⅜-½	–$1143
> ½-¾	–$2095
> ¾-1	–$2667

1 inch = 25.4 mm, 1 ft = 0.305 m

These factors are applied according to the units listed in the tables.

QC/QA Tests for Wisconsin

Table 153. Wisconsin QC/QA tests.

Test	QC Testing by Contractor	Verification Testing by Wisconsin DOT	Acceptance by Wisconsin DOT
Material passing #200	Two per day	NA	NA
Aggregate gradation	0-1000 tons: One per day 1000-2000 tons: Two per day 2000+ tons: Three per day	NA	NA
Slump	One per sublot	NA	NA
Aggregate moisture	Two per day	NA	NA
Air content	One per sublot	One per lot	Based on QC test
Profilograph	As soon as concrete has cured enough for testing	NA	Based on QC test
Thickness	Two per 250 ft	Two per day	Based on QC test
Compressive strength	Three cylinders per sublot	One per five lots	Based on QC test
Temperature	One per sublot	NA	NA

NA = Not applicable
1 ft = 0.305 m, 1 ton = 0.907 Mg

APPENDIX E: MINUTES FROM THE FIRST PANEL MEETING

Minutes of the March 1, 1999, Meeting of the Pooled Fund States for Optimal Acceptance Procedures for Statistical Specifications

Present:	Chris Abadie	Louisiana
	Roger Apple	Pennsylvania
	Ataur Bacchus	Ontario
	Jim Burati	Clemson University (project principal investigator)
	Steve DeWitt	North Carolina
	Doug Dirks	Illinois
	Steve Gage	Connecticut
	Chuck Hughes	Consultant (project consultant)
	Kurt Johnson	Wisconsin
	Peter Kopac	FHWA
	Rick Kreider	Kansas
	Bill Maupin	Virginia
	David Miller	Minnesota
	Tom Reis	Iowa
	Deniz Sandhu	New York
	Jeff Seiders	Texas
	Chris Williams	FHWA
Unable to Attend:	Milt Fletcher	South Carolina
	Rudy Malfabon	Washington
	Gary Selmi	Nevada
	Al Stanley	Idaho
	Ken Stoneman	Oregon
	Richard Weed	New Jersey

(1) The meeting began at 8:10 a.m. with a few introductory comments by Peter Kopac.

(2) Jim Burati then went through some administrative issues, including the agenda for the meeting (see attachment 1), participant introductions, and travel reimbursement procedures.

(3) Jim Burati then briefly summarized the progress of the project to date, including the literature search, the summary of State specifications, and the flowcharts of the specification development process.

Action Item: All participants are asked to review the specification summary information for their State and send any corrections or additions to Reagan Prince by fax (864-656-2670) or e-mail (jprince@clemson.edu).

(4) Chuck Hughes then reviewed the phase I flowchart of the specification development process. Comments and suggestions that were made included:

- Peter Kopac indicated that he thought there was a need to add an item, “Establish criteria for success,” in phase I, possibly as item 2.3.
- Jeff Seiders indicated that item 5, “Confirm interest and commitment,” was particularly important.
- Ataur Bacchus indicated that an executive summary document should be available for review throughout the specification development process. He also indicated that there were three objectives for the specification development process:
  1. Obtain effective specifications.
  2. Be able to explain the specifications at the practitioner level.
  3. Be able to convince executives (even if management changes) of the benefits of the specification.

Action Item: Jim Burati will revise the phase I flowchart to incorporate “Establish criteria for success” as an item. Chuck Hughes will then make any corresponding changes to the text that accompanies the flowchart.

(5) Chuck Hughes then reviewed the phase II flowchart of the specification development process. Comments and suggestions that were made included:

- Jeff Seiders indicated that it is necessary to retain method provisions and stipulated engineering practices for cases where good test procedures or measures are not available.
- Rick Kreider asked in how much detail “Dispute Resolution” would be covered in the manual. It was indicated that only the general concepts would be presented.
- It was decided that an effort should be made to ensure that the terminology in the flowcharts and the manual should be consistent with the Code of Federal Regulations, 23 CFR Part 637.
- Steve DeWitt asked whether it would be possible to put all of the phase II flowchart on a single page. Options will be explored to try to make it easier to follow the flow through the chart.
- Chris Abadie suggested that flowchart items similar to boxes 20 through 25 should be incorporated into the QC section (i.e., boxes 11 through 13).
- It was suggested that a decision diamond should be incorporated after box 25, “Obtain data,” to check whether the data warrant use of the property for acceptance purposes.

Action Item: Jim Burati will investigate revising the phase II flowchart to incorporate the items suggested in the bullets above. Chuck Hughes will then make any corresponding changes to the text that accompanies the flowchart.

(6) Chuck Hughes then reviewed the phase III flowchart of the specification development process. Comments and suggestions that were made included:

- Peter Kopac pointed out that it would be necessary to modify phase III to be consistent with the changes in phase I.

Action Item: Jim Burati will revise the phase III flowchart to ensure that it is consistent with phase I. Chuck Hughes will then make any corresponding changes to the text that accompanies the flowchart.

(7) Jim Burati then reviewed the preliminary specification analyses that have been conducted. These included:

- Bias in OC curves for PWL/PD.
- Precision in OC curves for PWL/PD.
- Precision of average project PWL/PD.
- Precision of individual payment estimates.
- Effect of bimodal distributions.
- Effect of skewed distributions.
- OC/expected payment curves for AAD plans.
- Comparison of expected payment between AAD and PWL plans.

In the interest of clarity, it was suggested that either PWL or PD, but not both, be used in the analyses that are done.

Action Item: The analyses that are performed on the project will all be reported in terms of PWL.

(8) Jim Burati then presented some additional analysis topics that might be considered during the project. These included:

- Use of Bayesian procedures for acceptance.
- Determination of the lot pay factors for individual properties.
- Determination of the composite pay factor for the lot.
- Validation/verification procedures for contractor/agency tests.
- Effect of non-normal populations.
- Effect of increased sampling and testing variability.
- Miscellaneous "bells and whistles," such as payment based on PWL, but with no price reductions if all individual tests are within the limits.

Action Item: Since there was not sufficient time to decide on the above items, it was decided that Jim Burati would distribute the items to the individual panel members for prioritizing and the items would then be returned to him for summarization.

(9) The meeting adjourned at 3:30 p.m.

Prepared and distributed by:

____

Jim Burati, Principal Investigator

APPENDIX F: ILLUSTRATIONS OF POSSIBLE RANGES FOR PD OR PWL ESTIMATES

The quality index, which is used to estimate PD (or PWL), is calculated from the sample mean and standard deviation. Let us assume that the sample standard deviation is constant. This is not true; however, it allows us to show a simplified illustration of the effect that the variability of the sample mean can have on the estimated PD (or PWL) value. The top plot in figure 98 shows a normal distribution (skewness coefficient = 0) with 10 percent below the lower specification limit.

The two solid-line plots in the figure illustrate the distribution of the sample means for a sample size of n = 3. The distribution of the sample means will be centered at the population mean, while the standard deviation of the distribution of the sample means (known as the *standard error*) will be equal to $\sigma_{\bar{X}} = \sqrt{\sigma / n}$, where σ is the population standard deviation. Therefore, the sample means will vary less than the original population from which they are drawn.

The distribution with the dashed line in the middle plot represents, for a sample size of n = 3, a normal distribution with the smallest mean that can be obtained from the original population (once again simplifying by saying that the normal distribution only extends 3 standard deviations on either side of the mean). It shows that it is possible to get an estimated PD value of greater than 50 (or a PWL of less than 50) even though the actual PD value for the population is 10 (the actual PWL is 90). The distribution with the dashed line in the bottom plot represents, for a sample size of n = 3, a normal distribution with the largest mean that can be obtained. It shows that it is possible to get an estimated PD value as small as zero (a PWL of 100) even though the actual PD value for the population is 10 (the PWL is 90).

These results are relatively consistent with the values shown in the bias histograms in chapter 6 and appendix G. The histograms in figure 32 indicate that a bias value of +90 was obtained. This indicates that a PD estimate of 100 was obtained during the simulation. The reason that the plot in figure 98 does not allow for a value this large is that, in developing figure 98, it was assumed that the standard deviation was constant. Therefore, the dashed-line distributions have the same spread as the original population. In reality, the sample standard deviation also varies and, therefore, it is possible to obtain a very low sample mean and, on the same lot, a very low sample standard deviation. This combination could result in an estimated distribution that lies entirely below the specification limit.

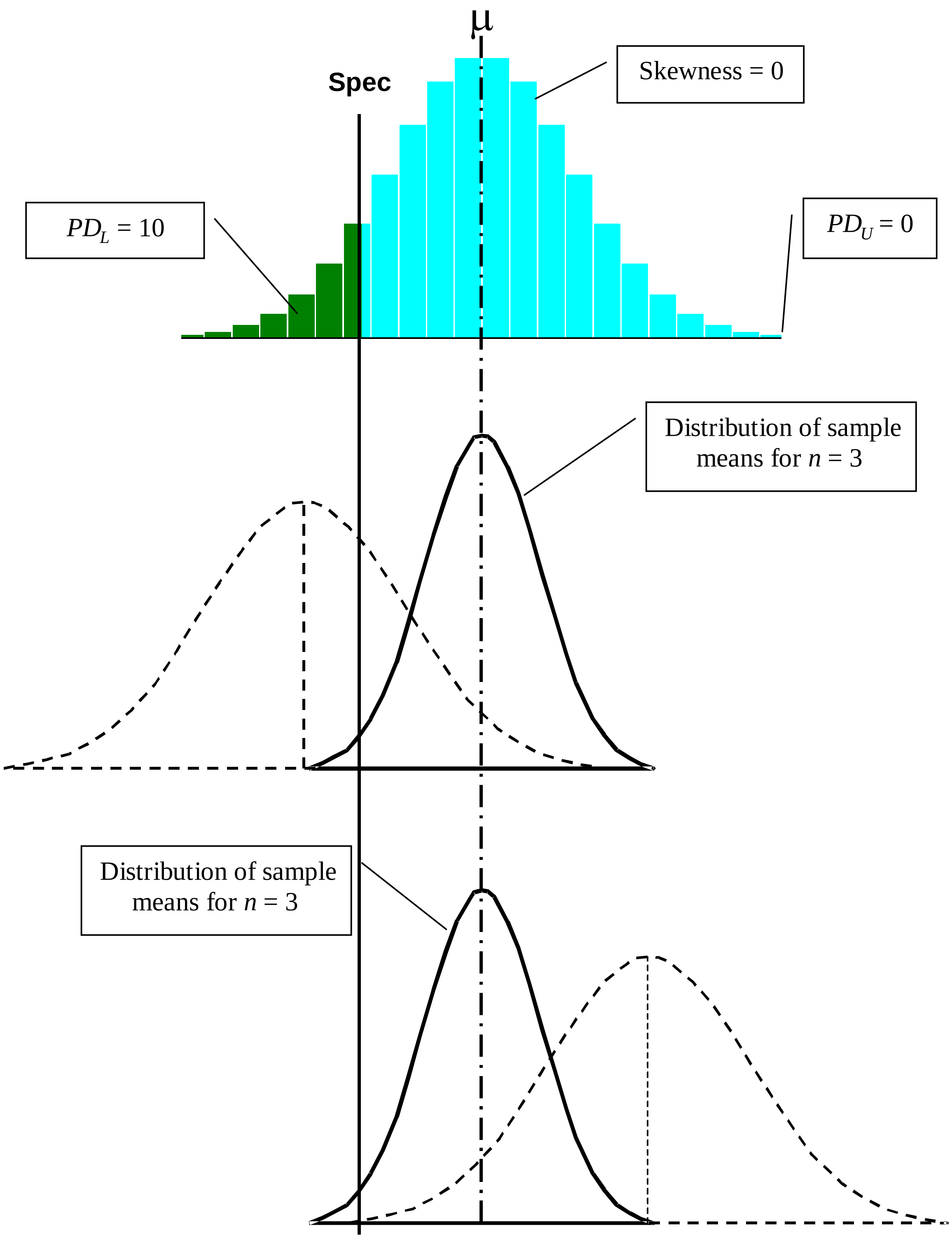


Figure 98. Spread of possible sample means for a normal distribution with 10 PD below the lower specification limit and sample size = 3.

The plots in figure 99 illustrate the same population as in figure 98, but with a sample size of $n = 10$. The larger sample size indicates that the distribution of the sample means will have less spread than that for a sample size of $n = 3$. It is still possible to have an estimated PD value of greater than 50; however, as drawn with the constant standard deviation, it would not be possible to have a PD estimate as low as zero. Once again, remember that these simplified examples assume that the standard deviation does not vary and remains equal to the population's standard deviation. They do, however, illustrate how natural sampling variability will lead to variability in the estimated PD (or PWL) values.

The remaining figures in appendix F illustrate different PD_L/PD_U divisions and different population PD values for the cases of symmetrical distributions (skewness = 0) and for skewed distributions (skewness coefficient = 1.0). They can be interpreted in the same way as the discussions for figures 98 and 99.

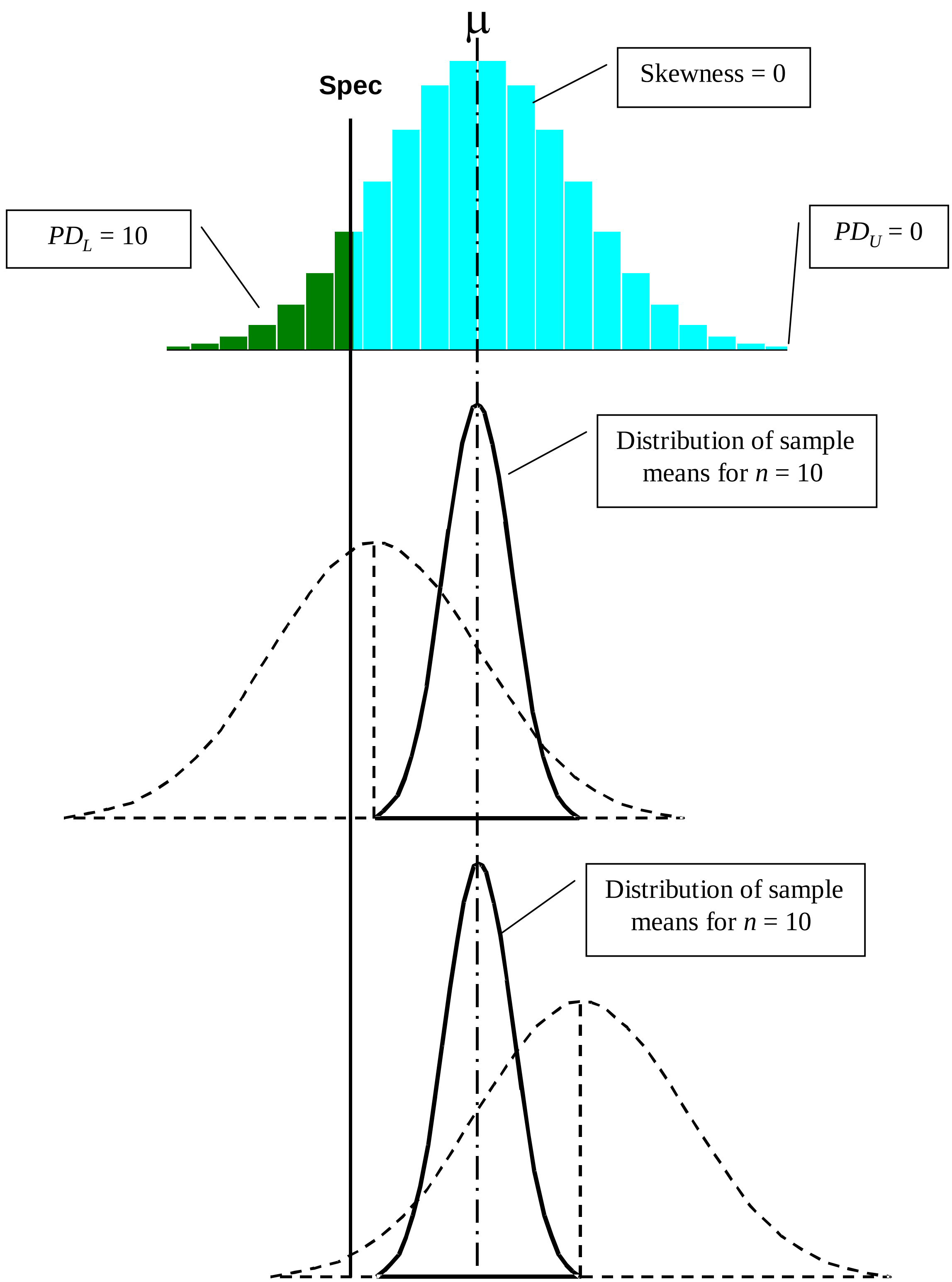


Figure 99. Spread of possible sample means for a normal distribution with 10 PD below the lower specification limit and sample size = 10.

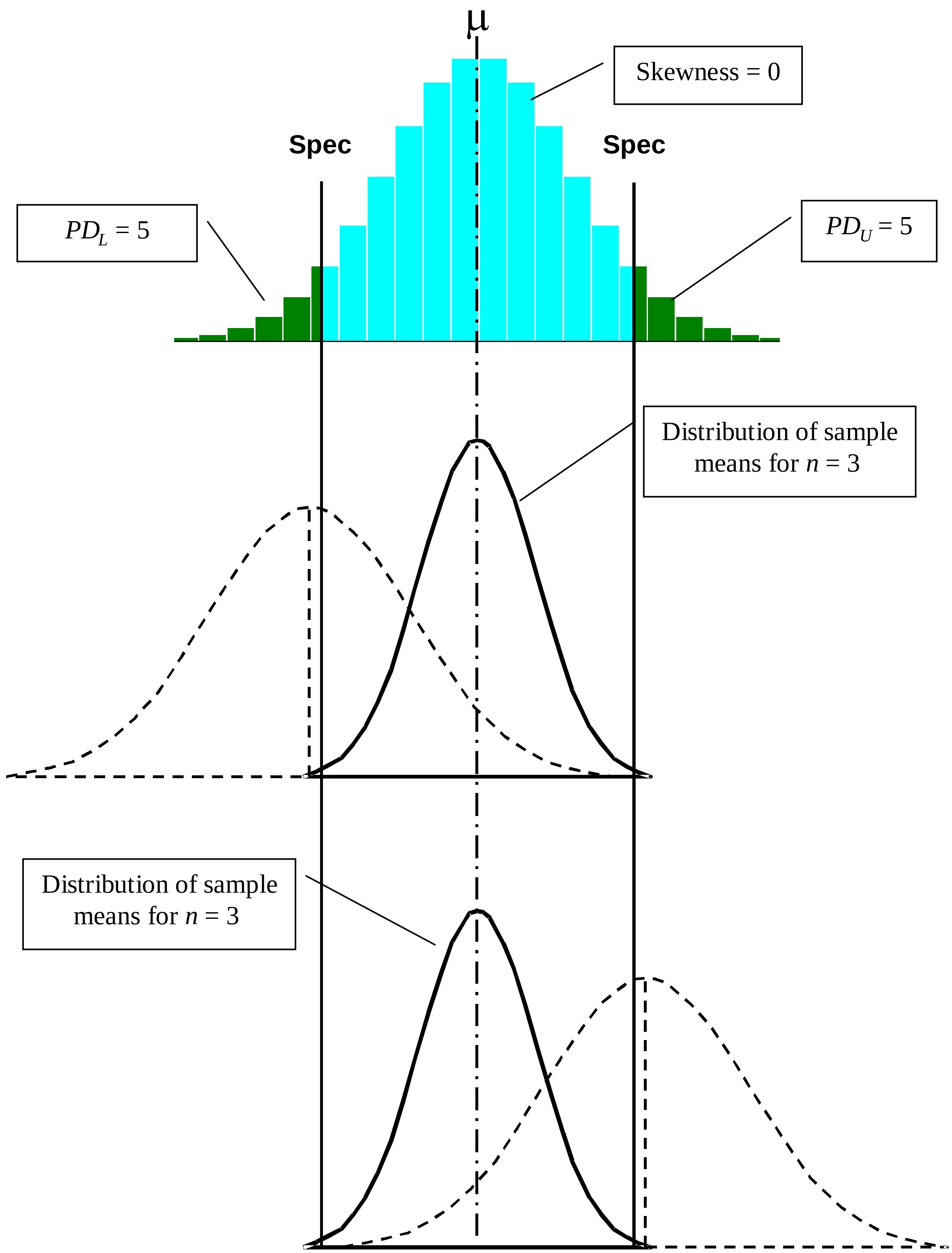


Figure 100. Spread of possible sample means for a normal distribution with 5 PD outside each specification limit and sample size = 3.

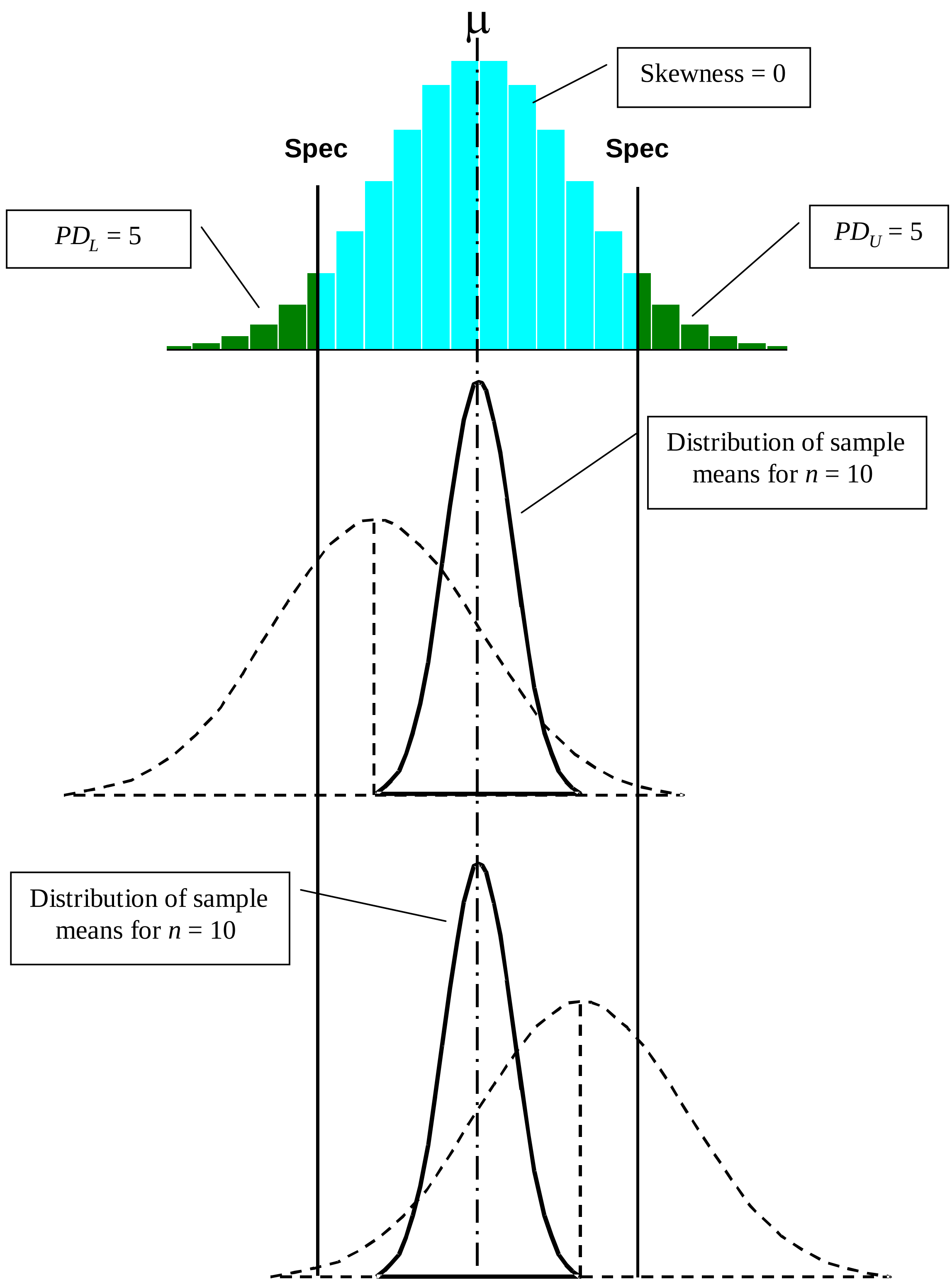


Figure 101. Spread of possible sample means for a normal distribution with 5 PD outside each specification limit and sample size = 10.

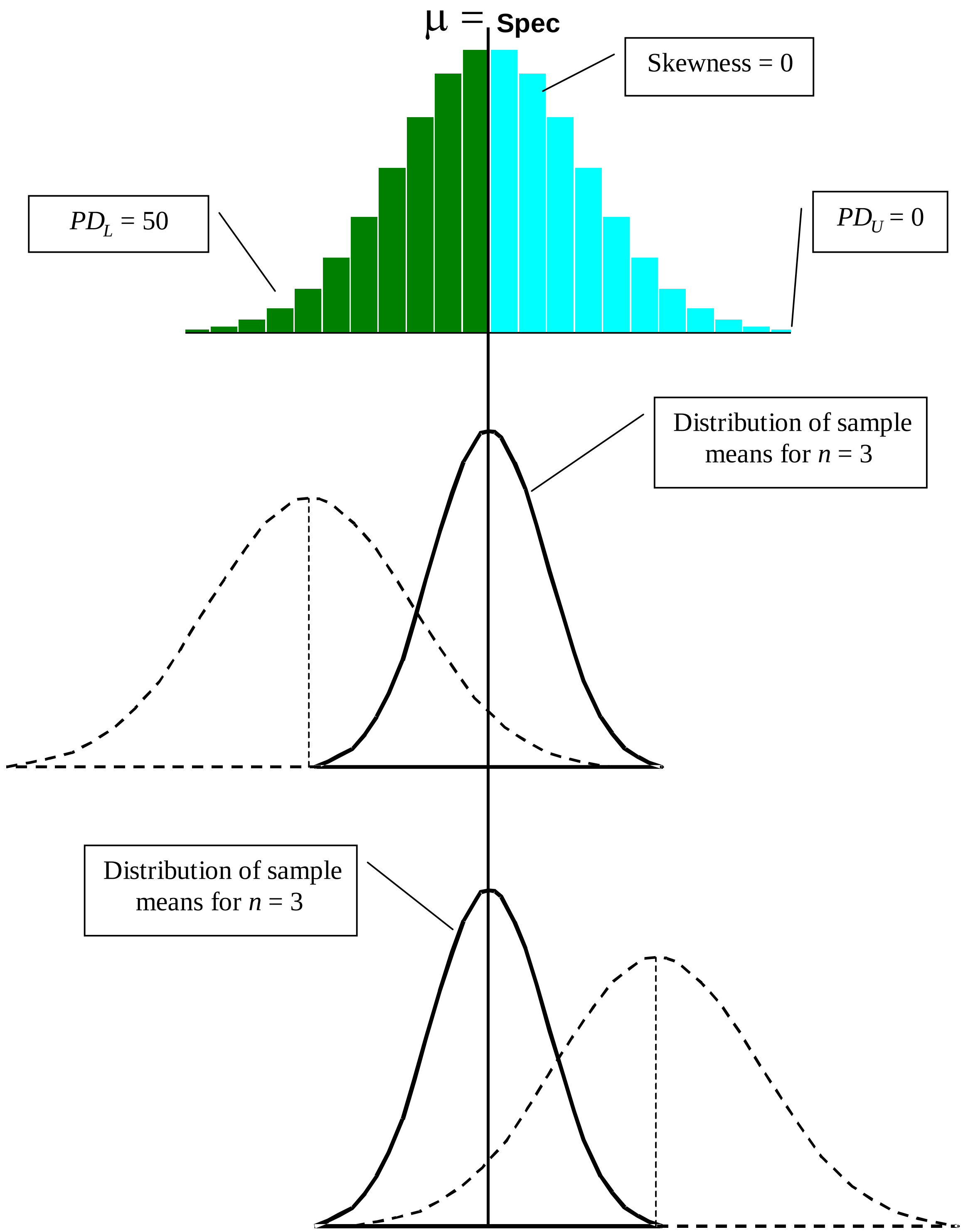


Figure 102. Spread of possible sample means for a normal distribution with 50 PD below the lower specification limit and sample size = 3.

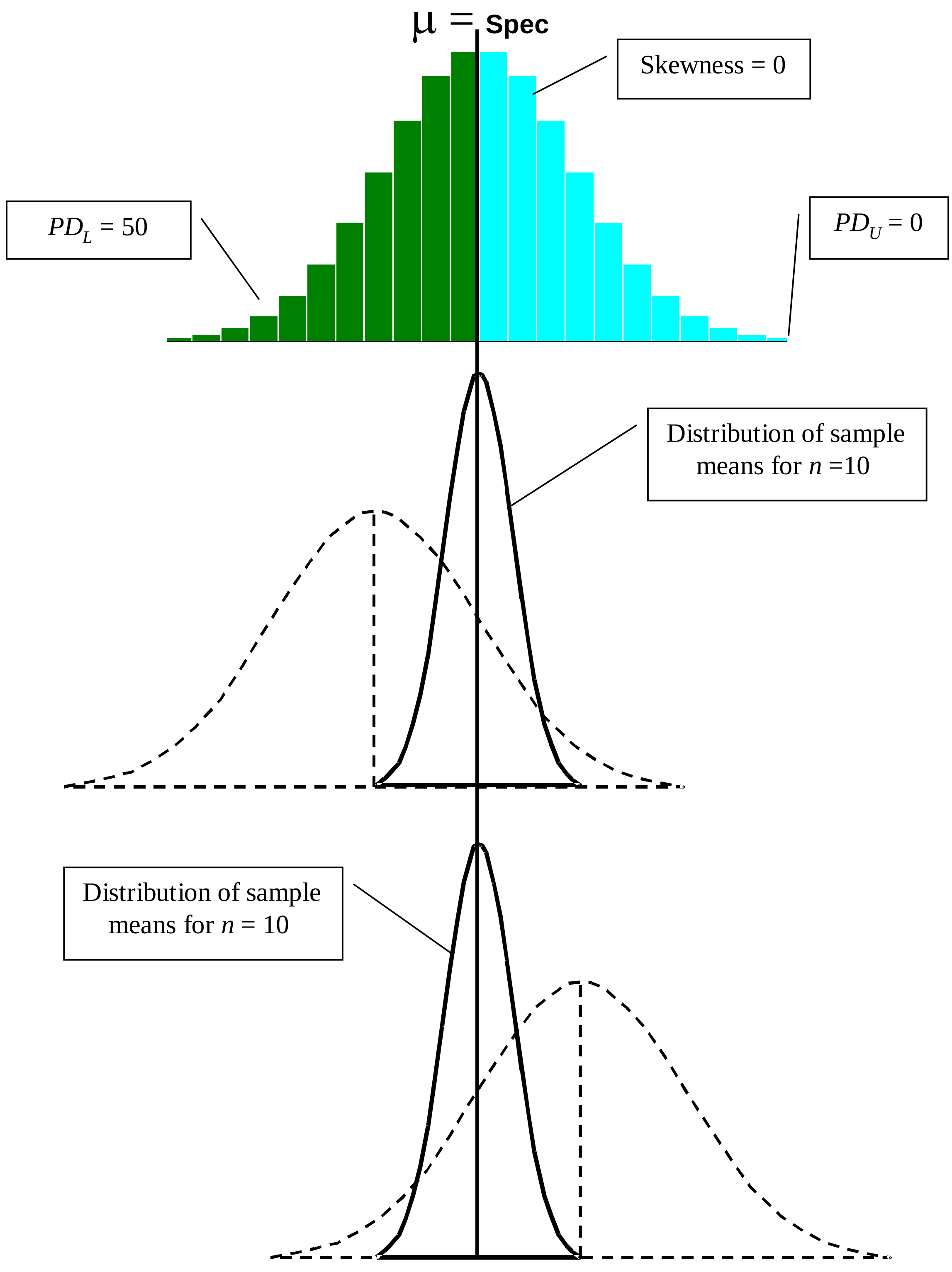


Figure 103. Spread of possible sample means for a normal distribution with 50 PD below the lower specification limit and sample size = 10.

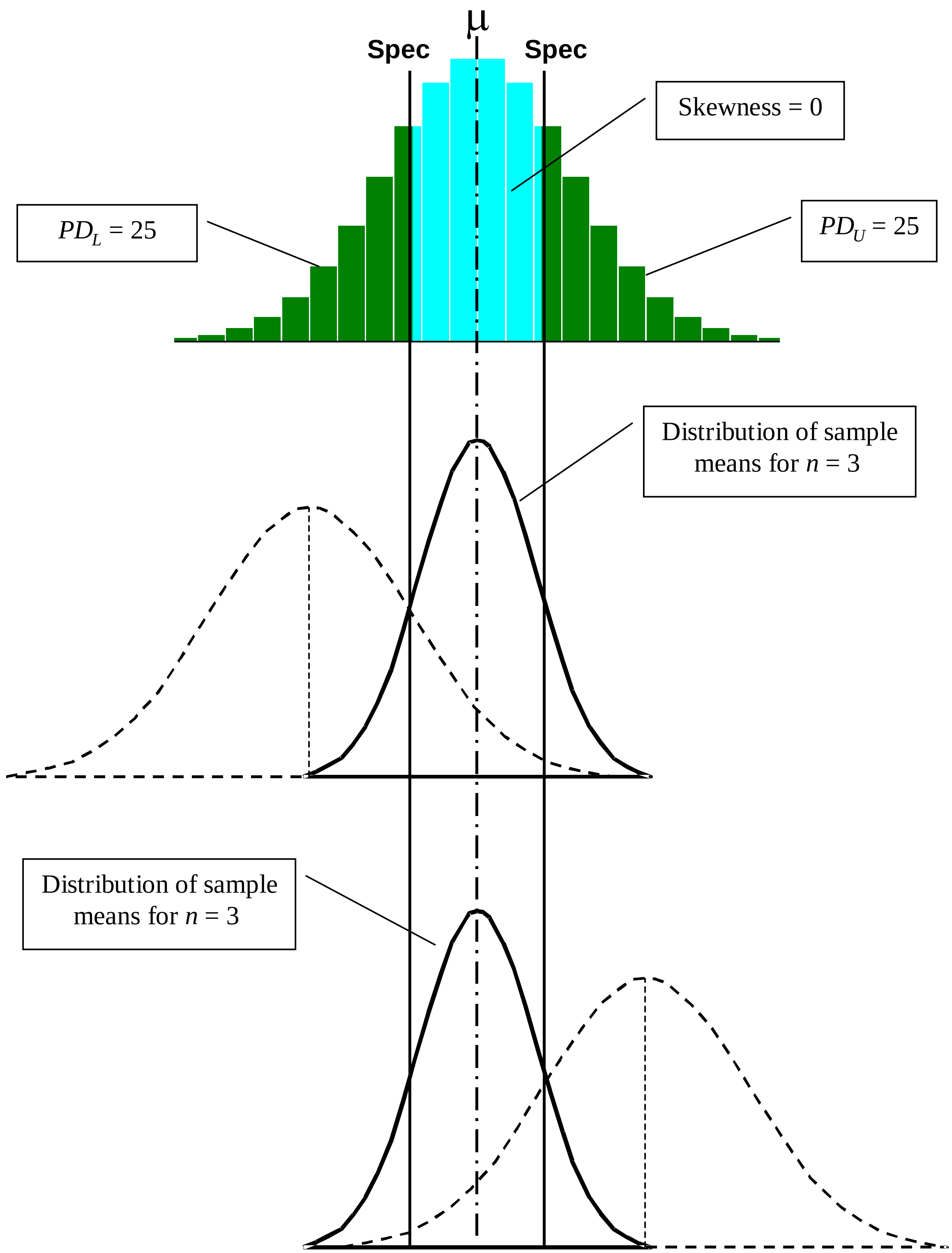


Figure 104. Spread of possible sample means for a normal distribution with 25 PD outside each specification limit and sample size = 3.

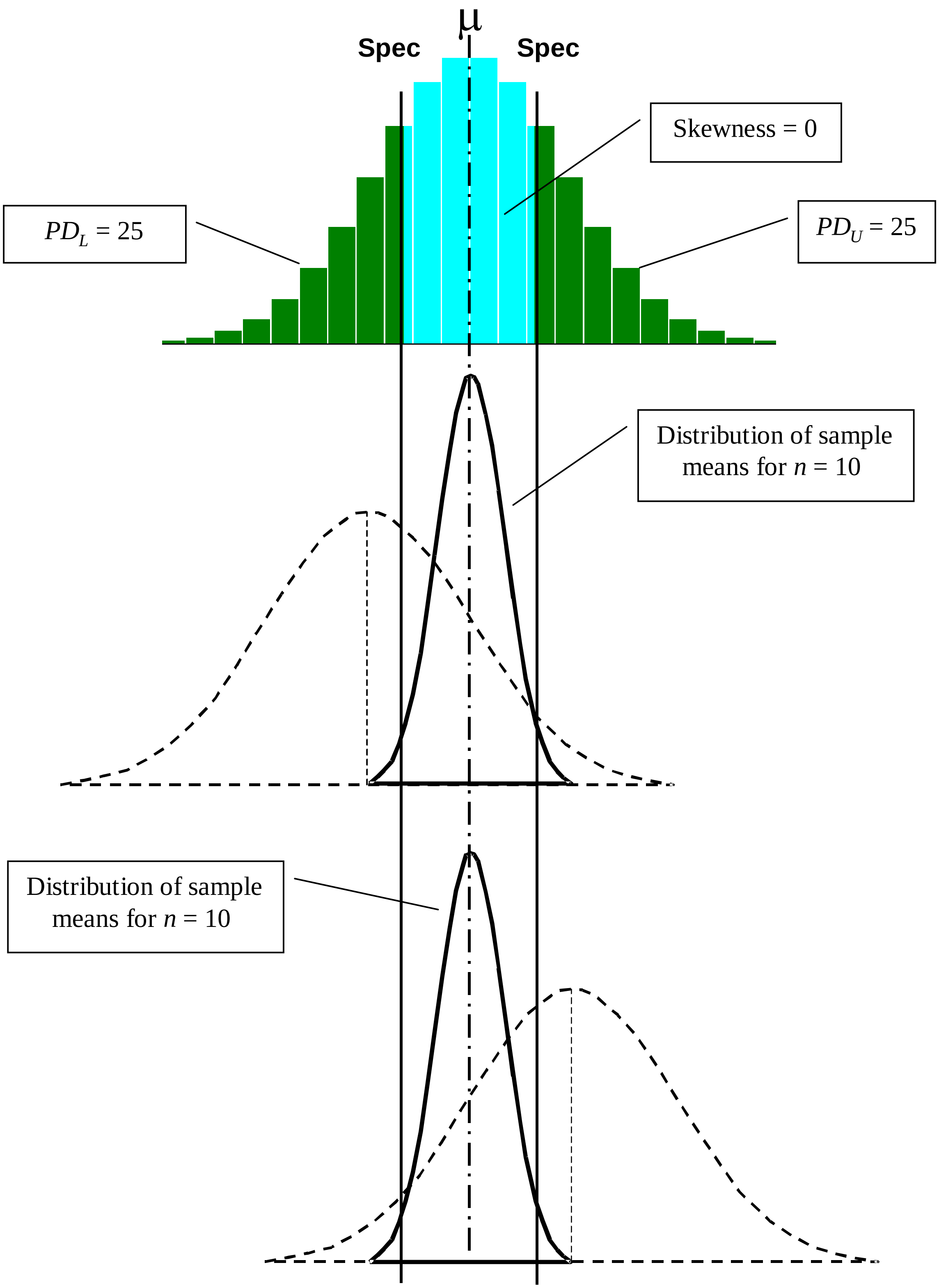


Figure 105. Spread of possible sample means for a normal distribution with 25 PD outside each specification limit and sample size = 10.

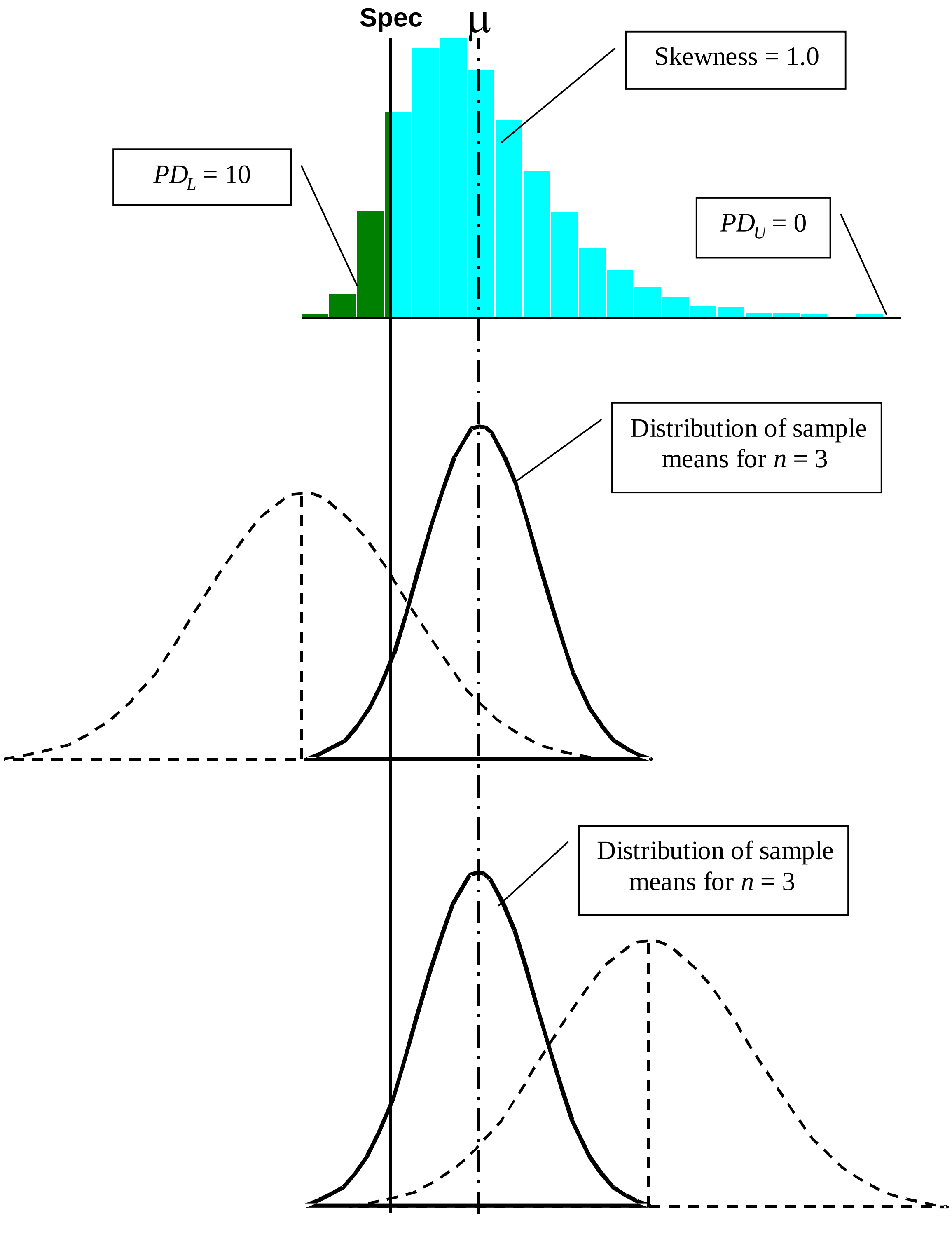


Figure 106. Spread of possible sample means for a distribution with skewness = 1.0, 10 PD below the lower specification limit, and sample size = 3.

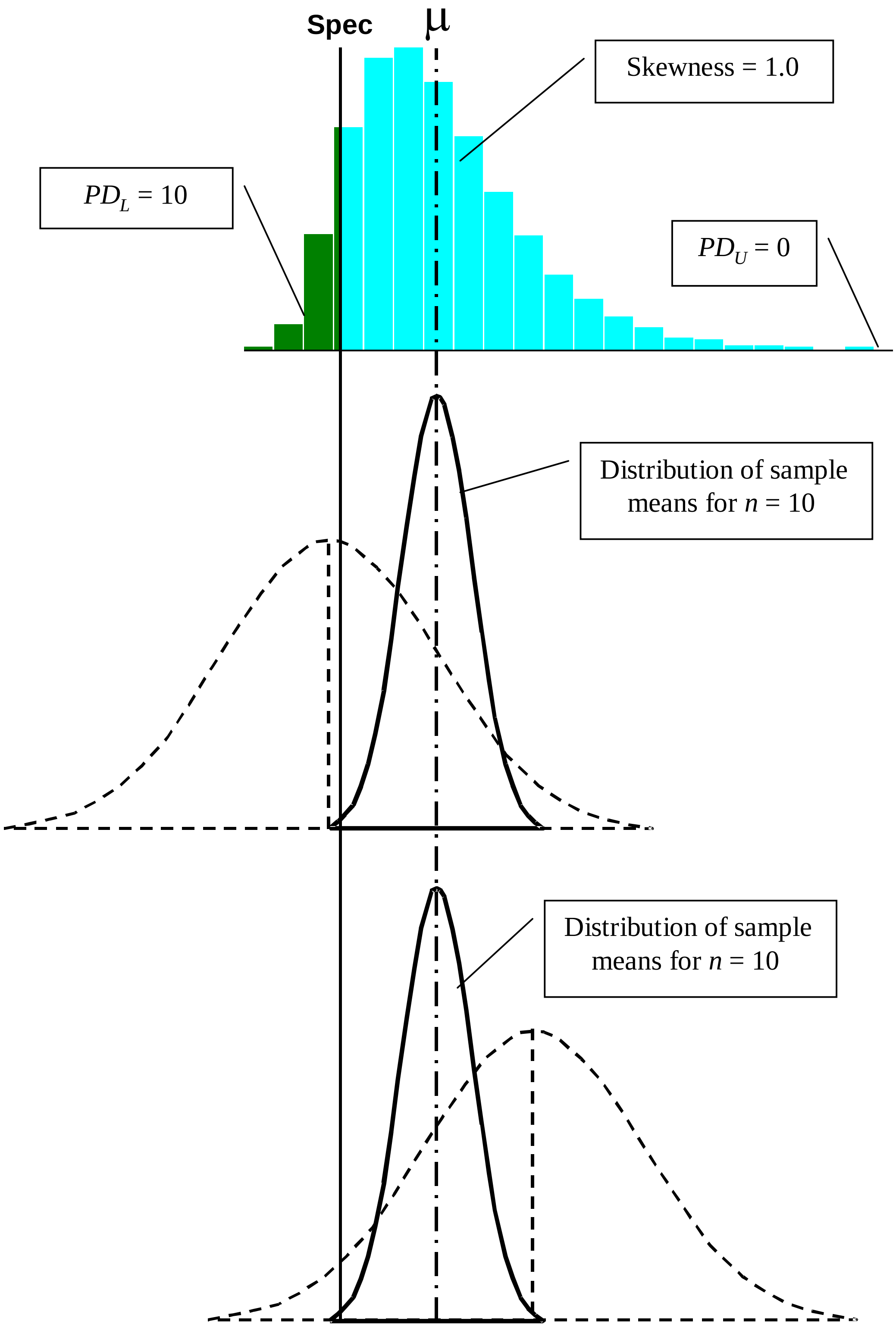


Figure 107. Spread of possible sample means for a distribution with skewness = 1.0, 10 PD below the lower specification limit, and sample size = 10.

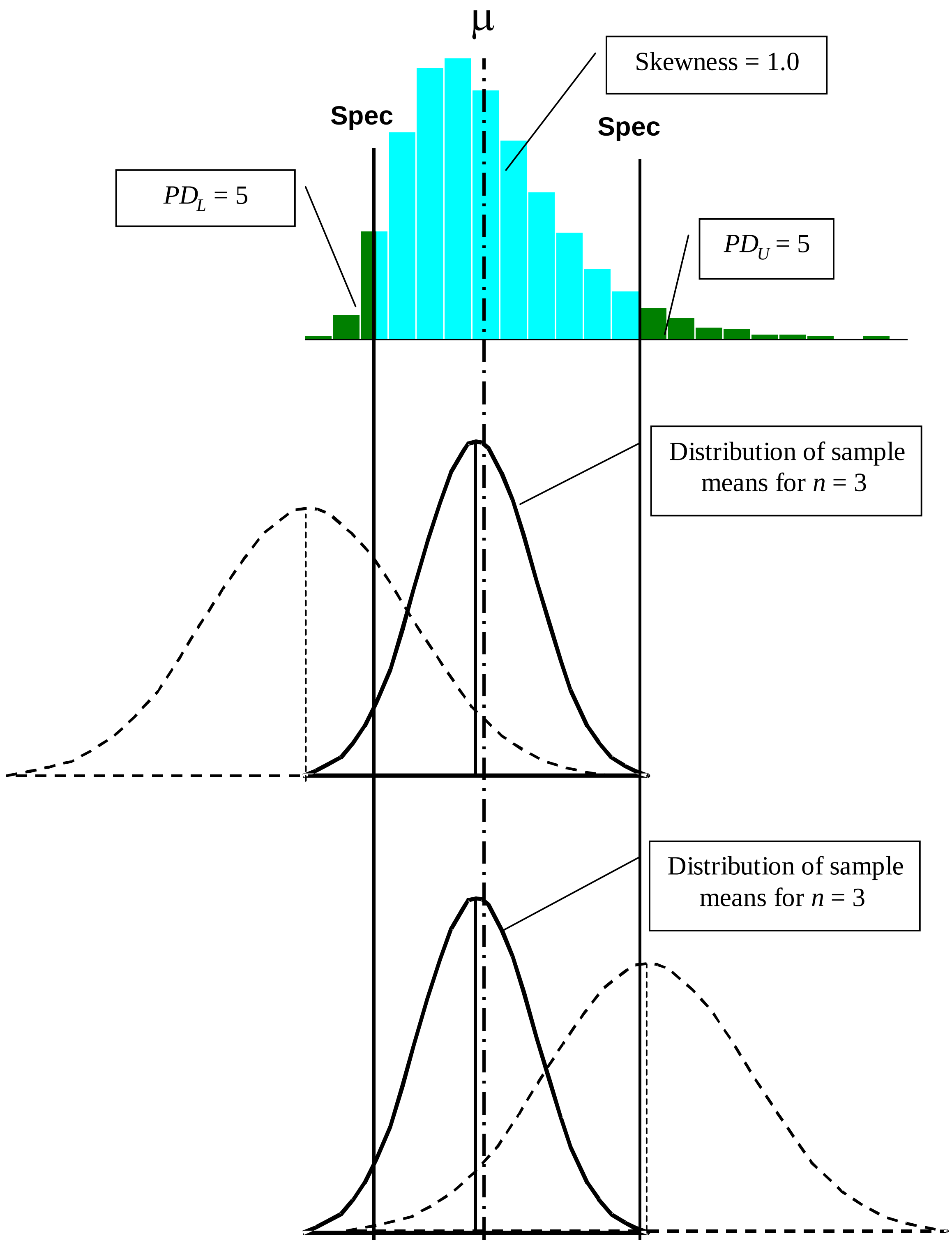


Figure 108. Spread of possible sample means for a distribution with skewness = 1.0, 5 PD outside each specification limit, and sample size = 3.

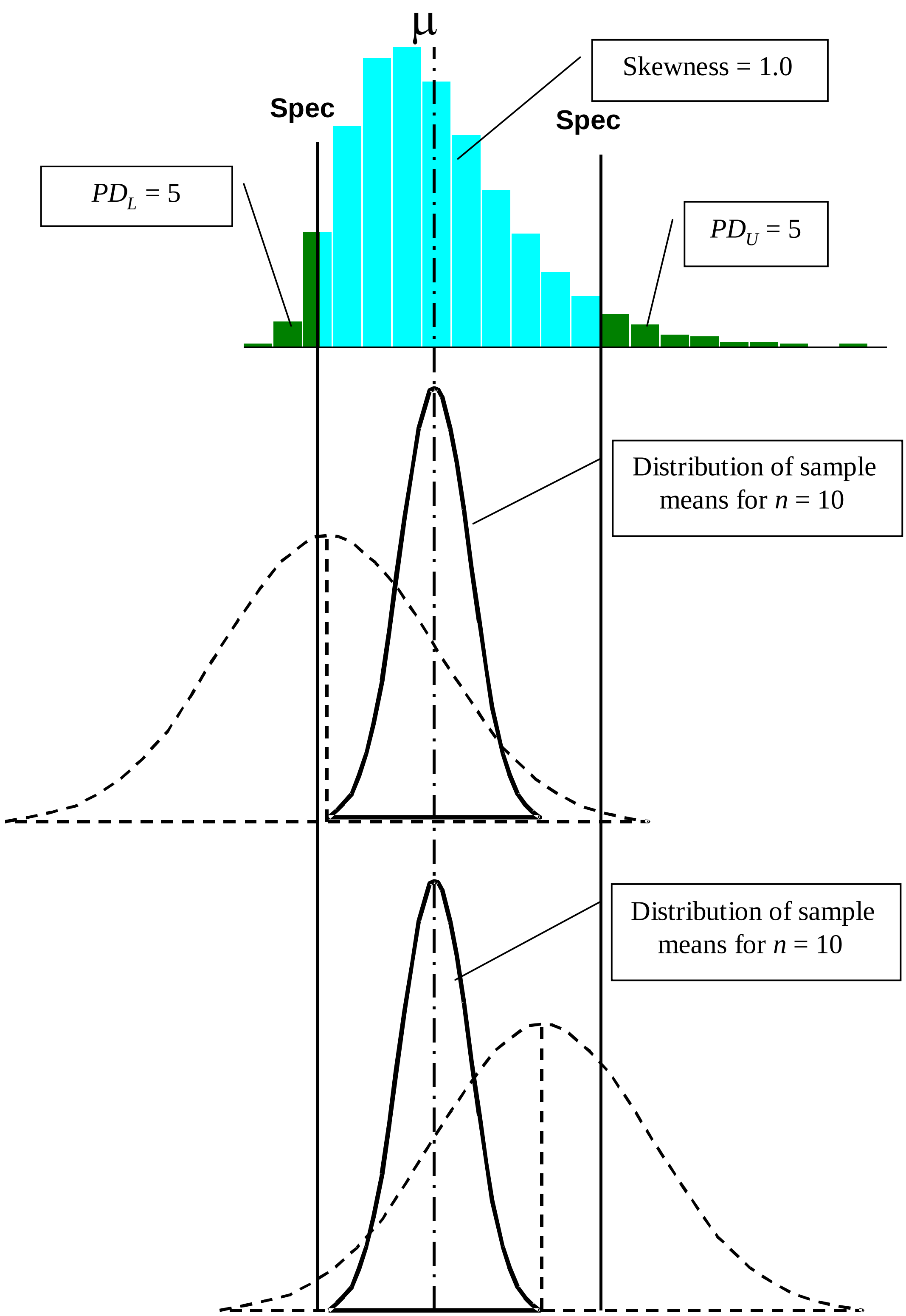


Figure 109. Spread of possible sample means for a distribution with skewness = 1.0, 5 PD outside each specification limit, and sample size = 10.

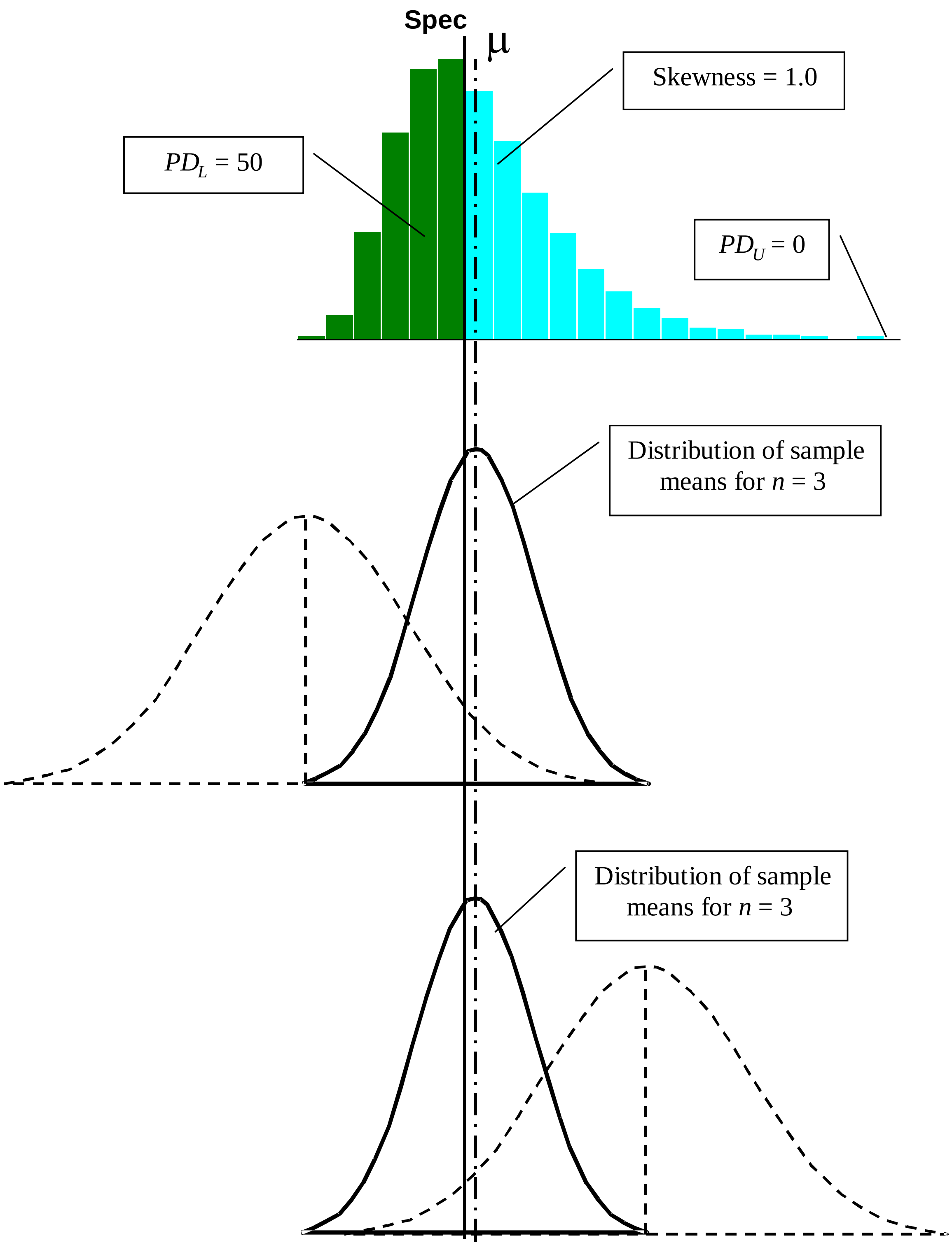


Figure 110. Spread of possible sample means for a distribution with skewness = 1.0, 50 PD below the lower specification limit, and sample size = 3.

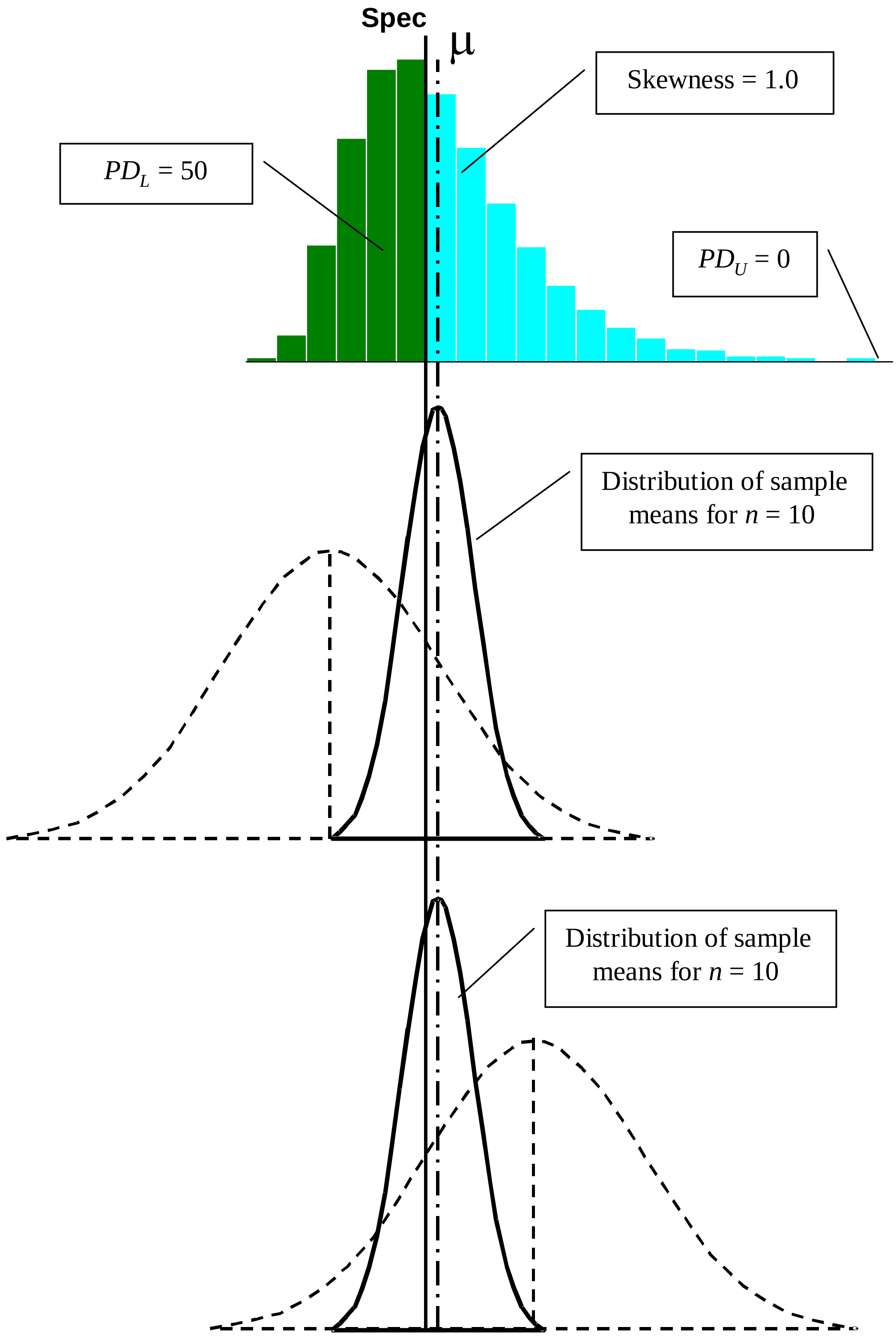


Figure 111. Spread of possible sample means for a distribution with skewness = 1.0, 50 PD below the lower specification limit, and sample size = 10.

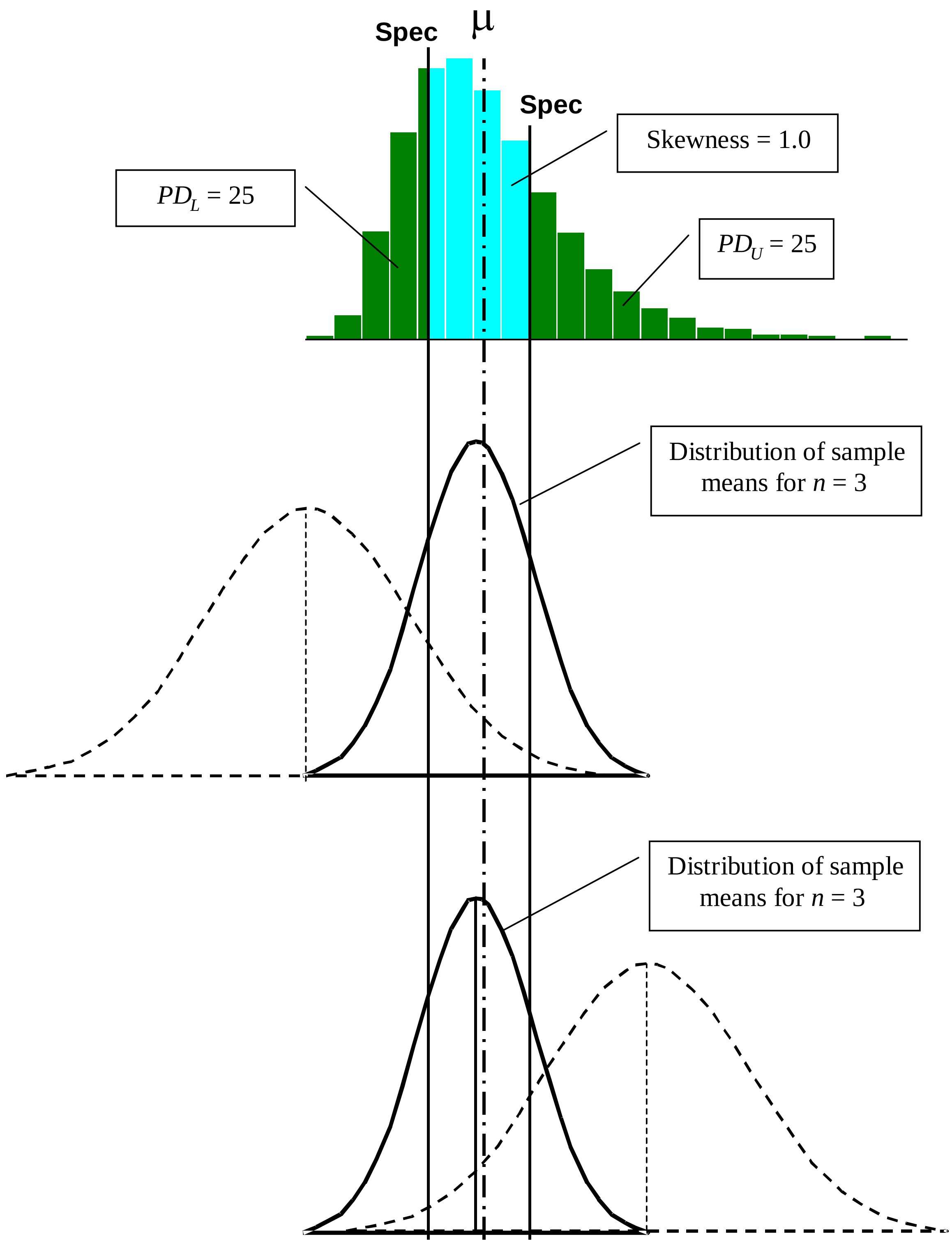


Figure 112. Spread of possible sample means for a distribution with skewness = 1.0, 25 PD outside each specification limit, and sample size = 3.

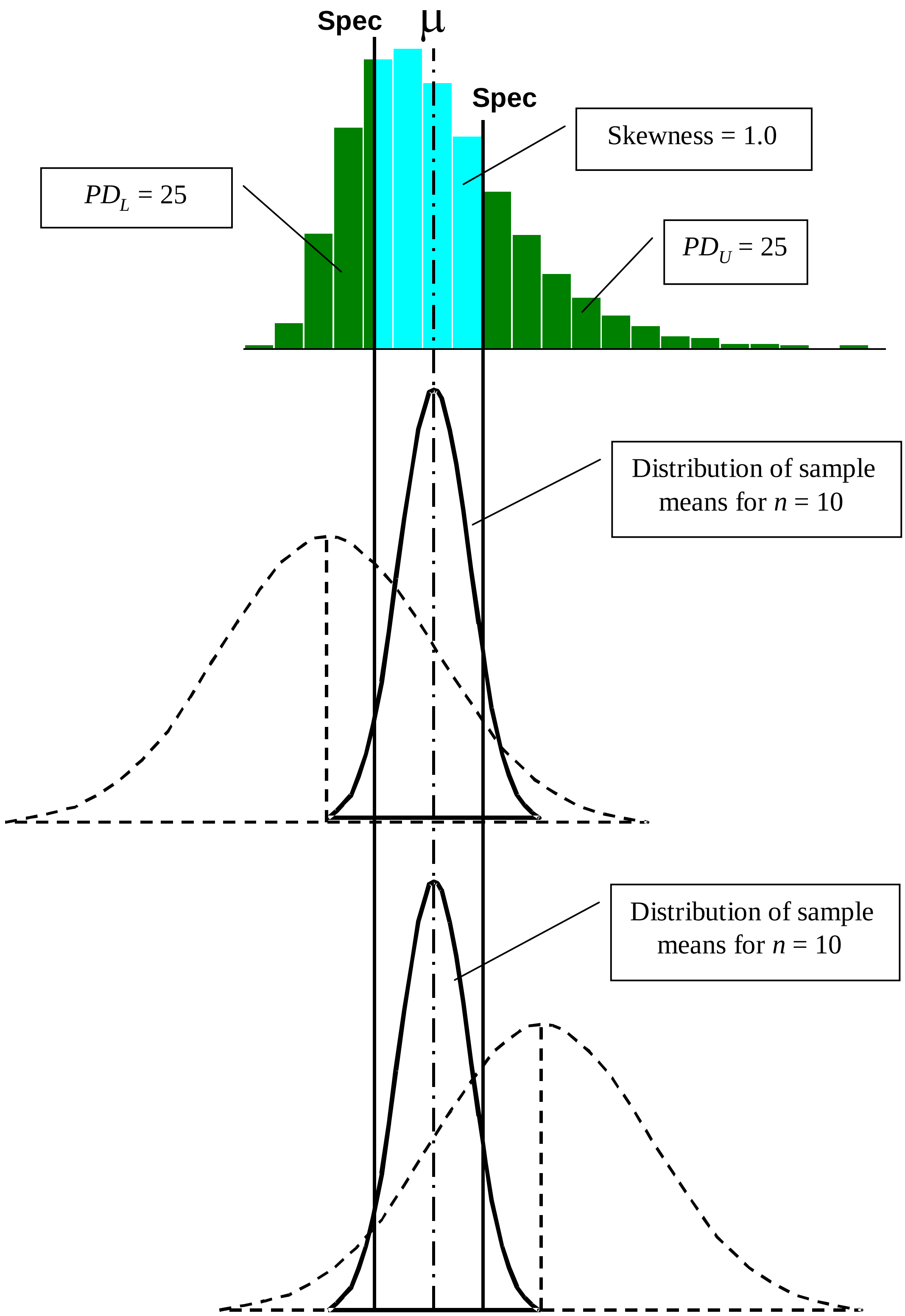


Figure 113. Spread of possible sample means for a distribution with skewness = 1.0, 25 PD outside each specification limit, and sample size = 10.

APPENDIX G: BIAS HISTOGRAMS FROM THE SIMULATION PROGRAM

The plots in appendix G illustrate the distributions of sample means for a number of different combinations of population PD, sample size, and skewness coefficients. These figures are an expansion of those shown and discussed in chapter 6.

Figure 114 shows how sample size can affect the variability of the PD estimates. This figure shows the complete program output for a symmetrical population (i.e., skewness coefficient = 0) that actually has 10 PD and for which the sample size is $n = 5$. The results of the nine divisions for PD_L/PD_U are clearly indicated. The BIAS/SE results indicate the average bias and the standard error (measure of variability of the PD estimates), respectively. These terms are discussed and explained in detail in chapter 6. The bias values are relatively closely distributed about zero, indicating that the method for estimating PD (and hence PWL) is not biased for a symmetrical distribution. This is to be expected since the PD estimating method assumes a symmetrical normal distribution.

Figure 115 shows portions of the program output for the same population, but with sample sizes of $n = 3$ and $n = 10$. Figures 116 and 117, and 118 and 119 show similar plots for populations with skewness coefficients = 1.0 and 2.0, respectively. The remaining figures illustrate other comparisons of the bias results. Figures 120 and 121 show the bias values histograms for sampling with the same sample size of $n = 5$ from symmetrical populations with different PD values (i.e., PD = 30 and PD = 50, respectively). Figures 122 and 123 show similar results for a sample size of $n = 10$. Figure 124 shows the bias values histograms for the same sample size of $n = 5$ and the same PD value = 50, but for populations with skewness coefficients = 0.0, 1.0, 2.0, and 3.0. Figures 125 through 127 show the complete program outputs for populations with a skewness coefficient = 3.0, a sample size of $n = 5$, and PD values = 10, 30, and 50. It is unlikely that populations with this degree of skewness will be found in practice.

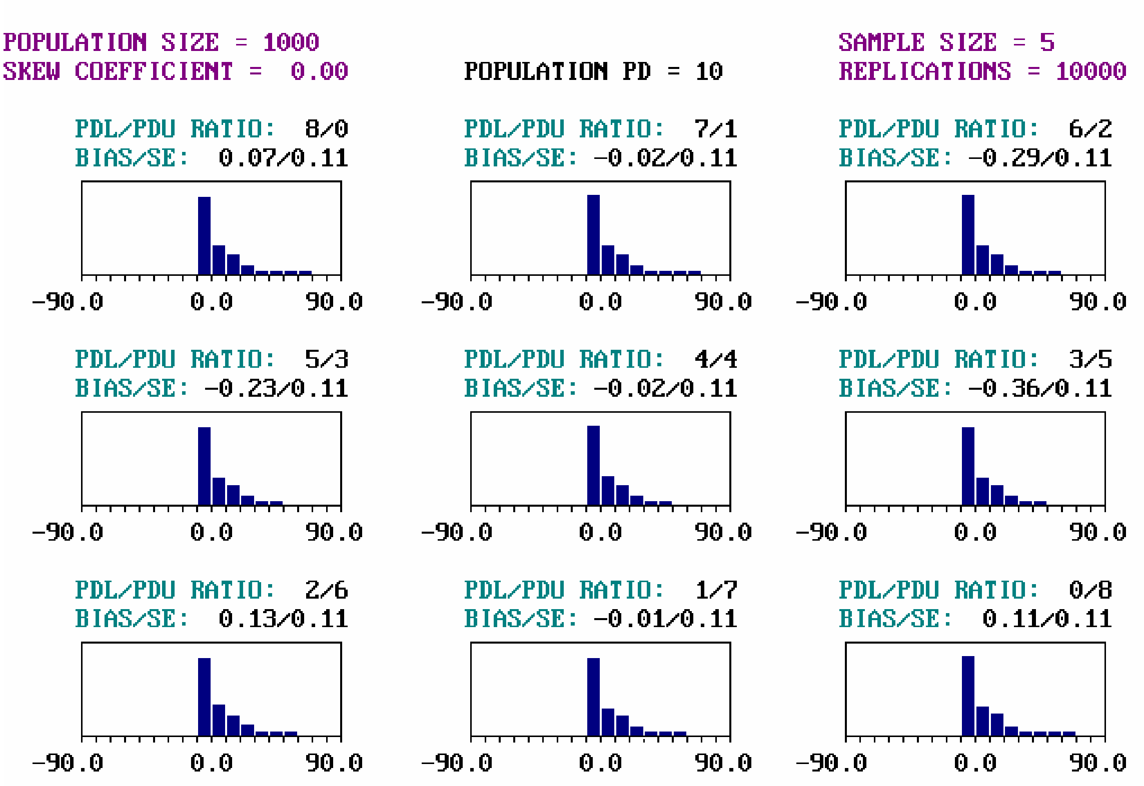


Figure 114. Sample output screen for a population with PD = 10, skewness coefficient = 0.00, and sample size = 5.

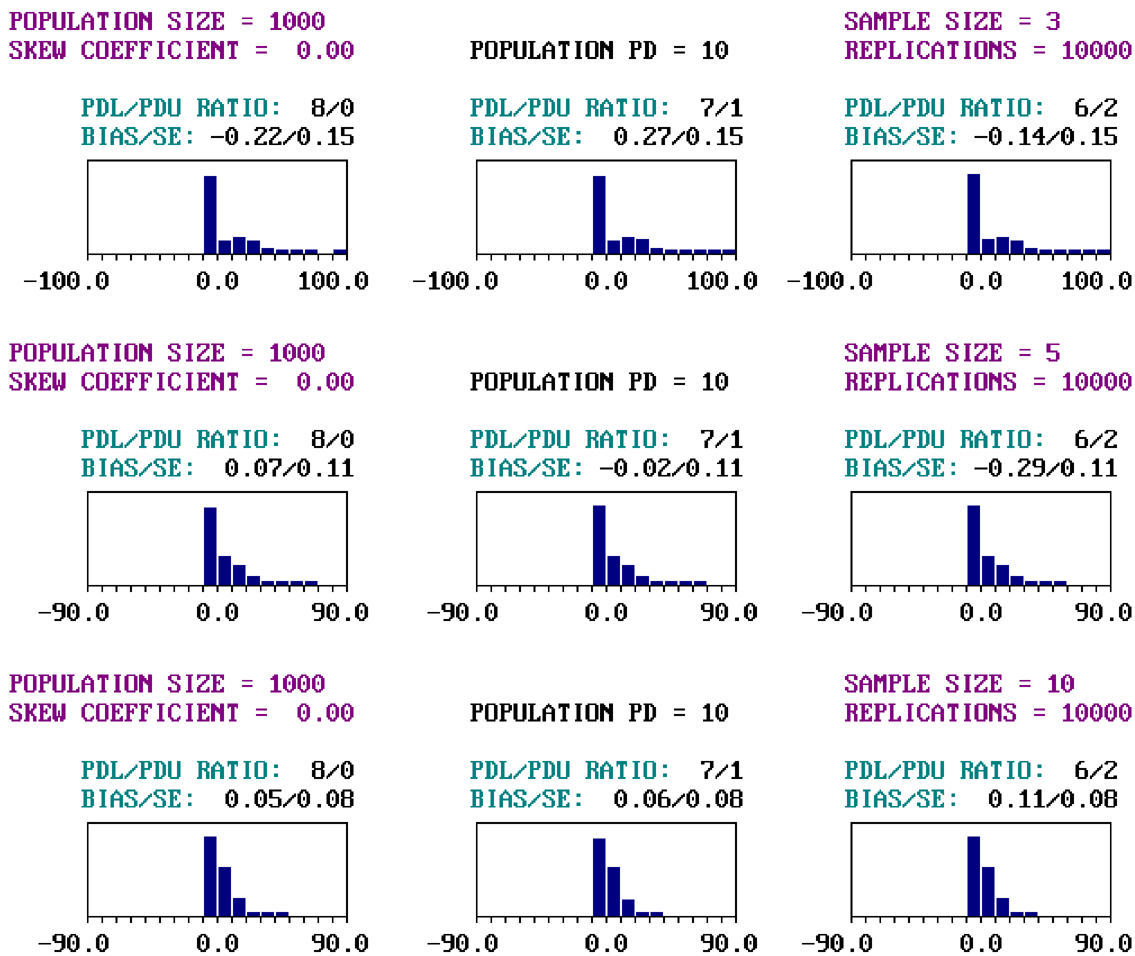


Figure 115. Portions of output screens for PD = 10, skewness coefficient = 0.00, and sample sizes = 3, 5, and 10.

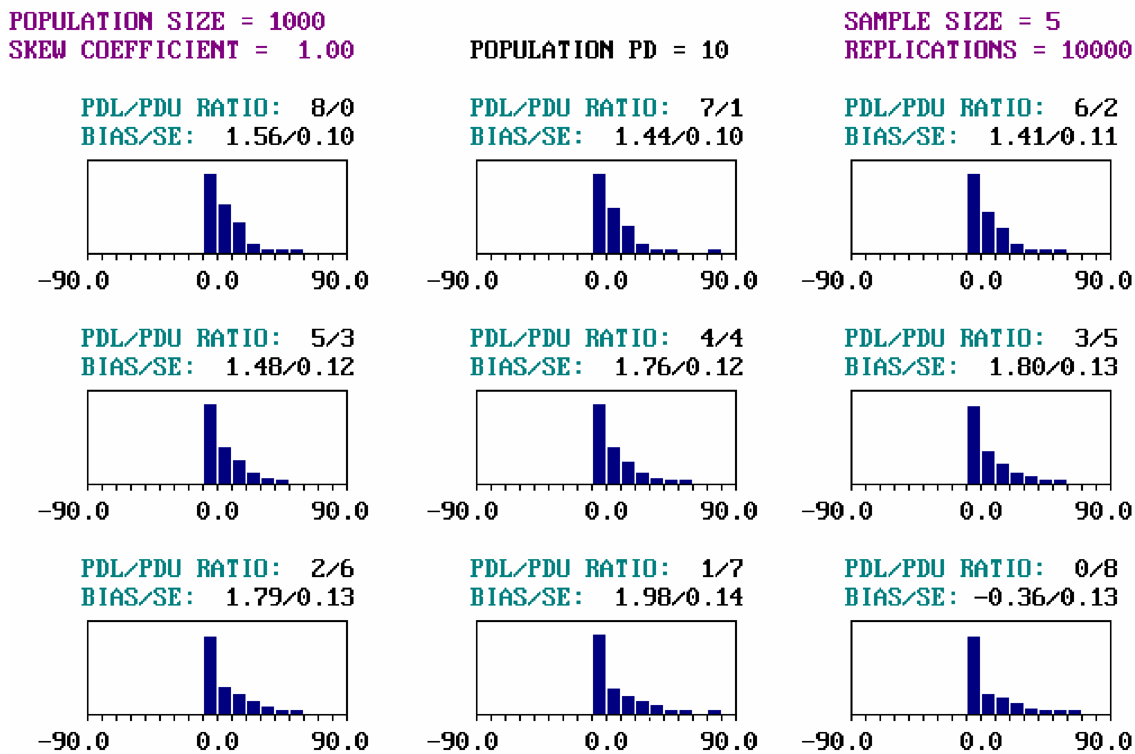


Figure 116. Sample output screen for a population with PD = 10, skewness coefficient = 1.00, and sample size = 5.

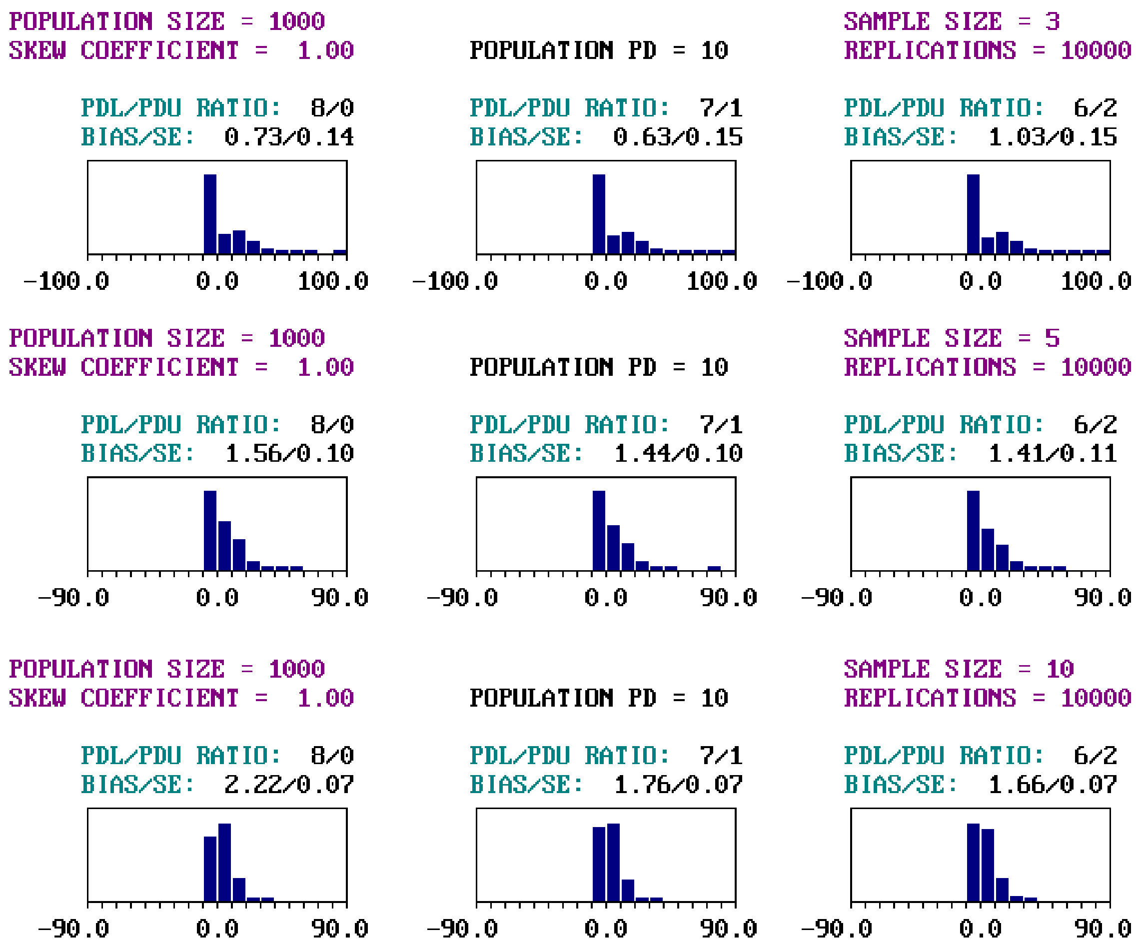


Figure 117. Portions of output screens for PD = 10, skewness coefficient = 1.00, and sample sizes = 3, 5, and 10.

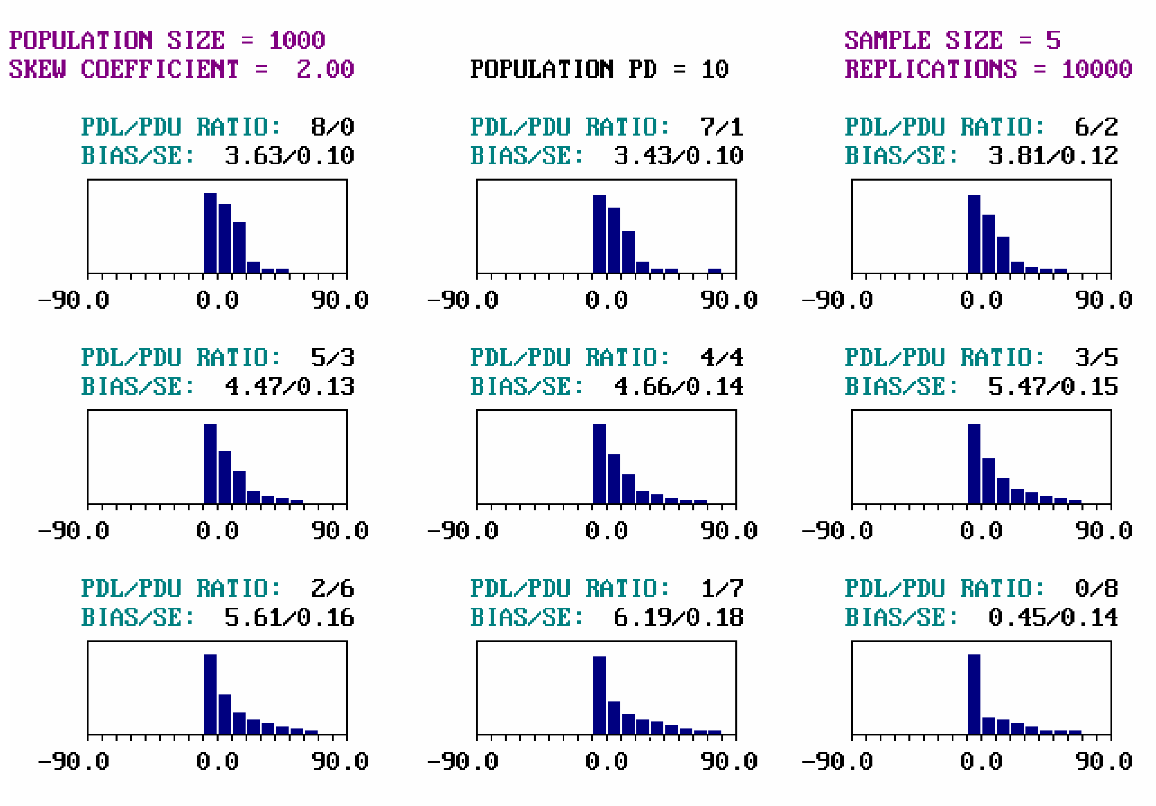


Figure 118. Sample output screen for a population with PD = 10, skewness coefficient = 2.00, and sample size = 5.

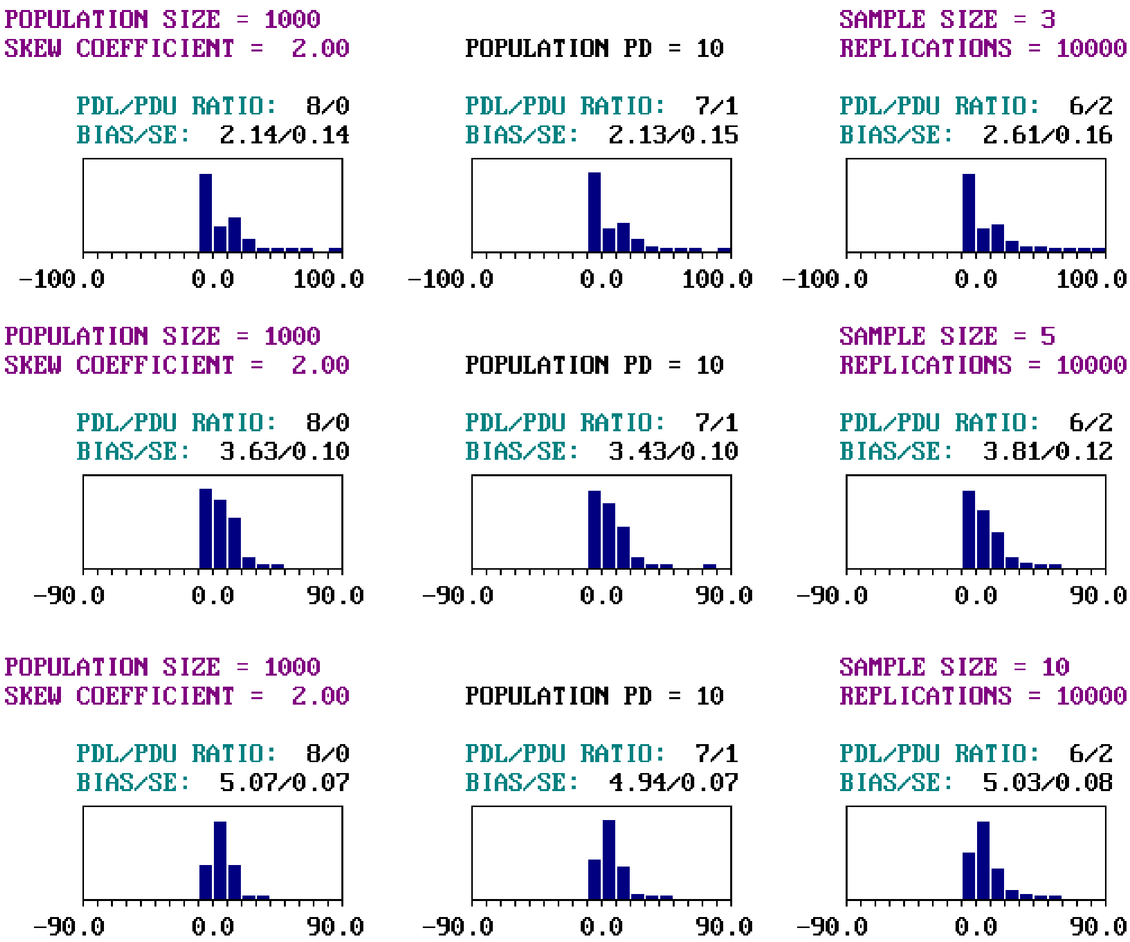


Figure 119. Portions of output screens for PD = 10, skewness coefficient = 2.00, and sample sizes = 3, 5, and 10.

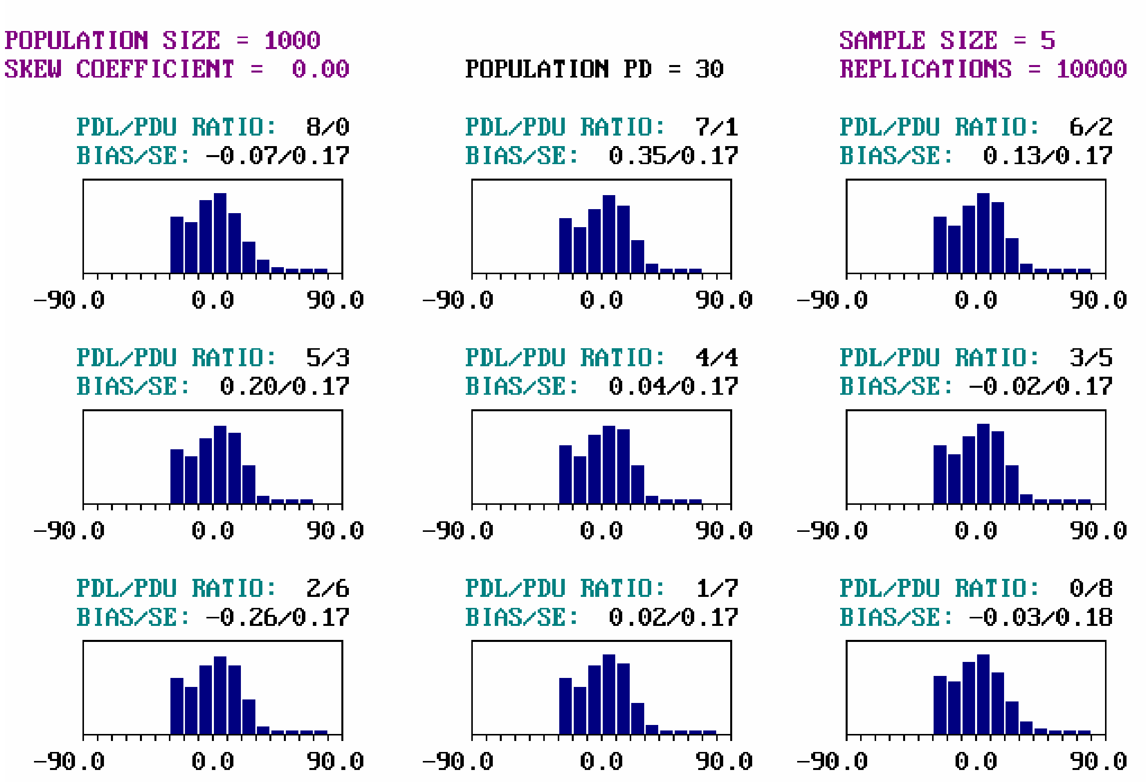


Figure 120. Sample output screen for a population with PD = 30, skewness coefficient = 0.00, and sample size = 5.

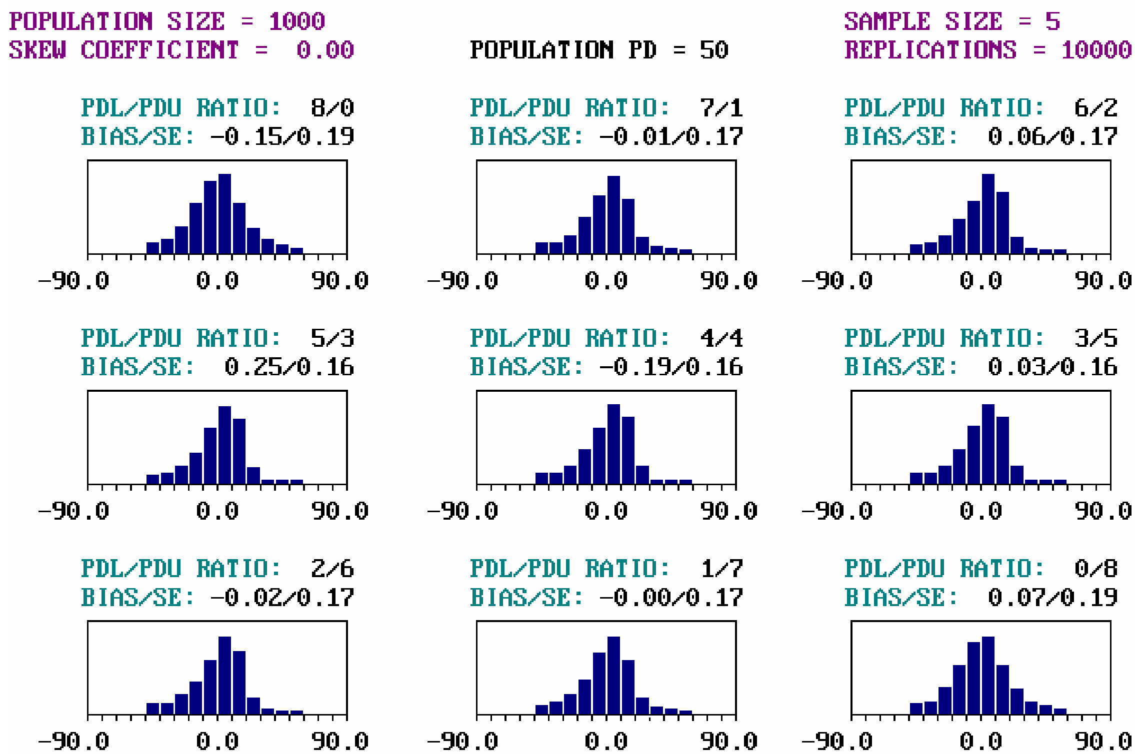


Figure 121. Sample output screen for a population with PD = 50, skewness coefficient = 0.00, and sample size = 5.

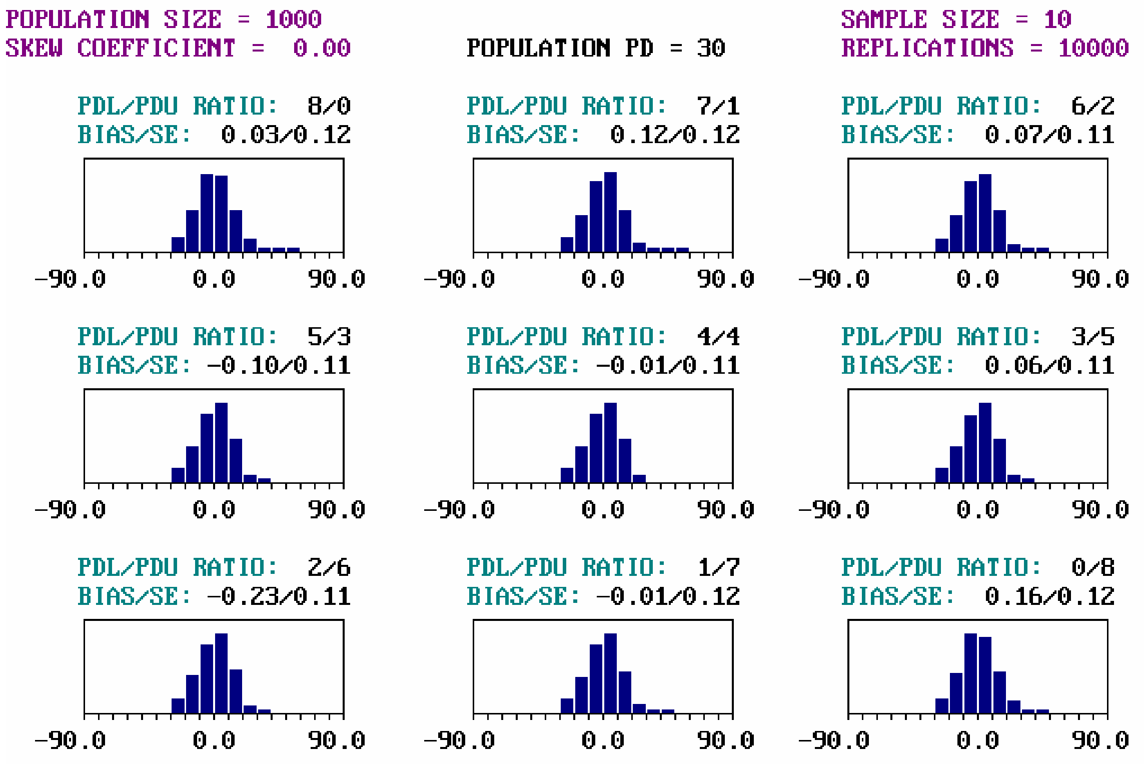


Figure 122. Sample output screen for a population with PD = 30, skewness coefficient = 0.00, and sample size = 10.

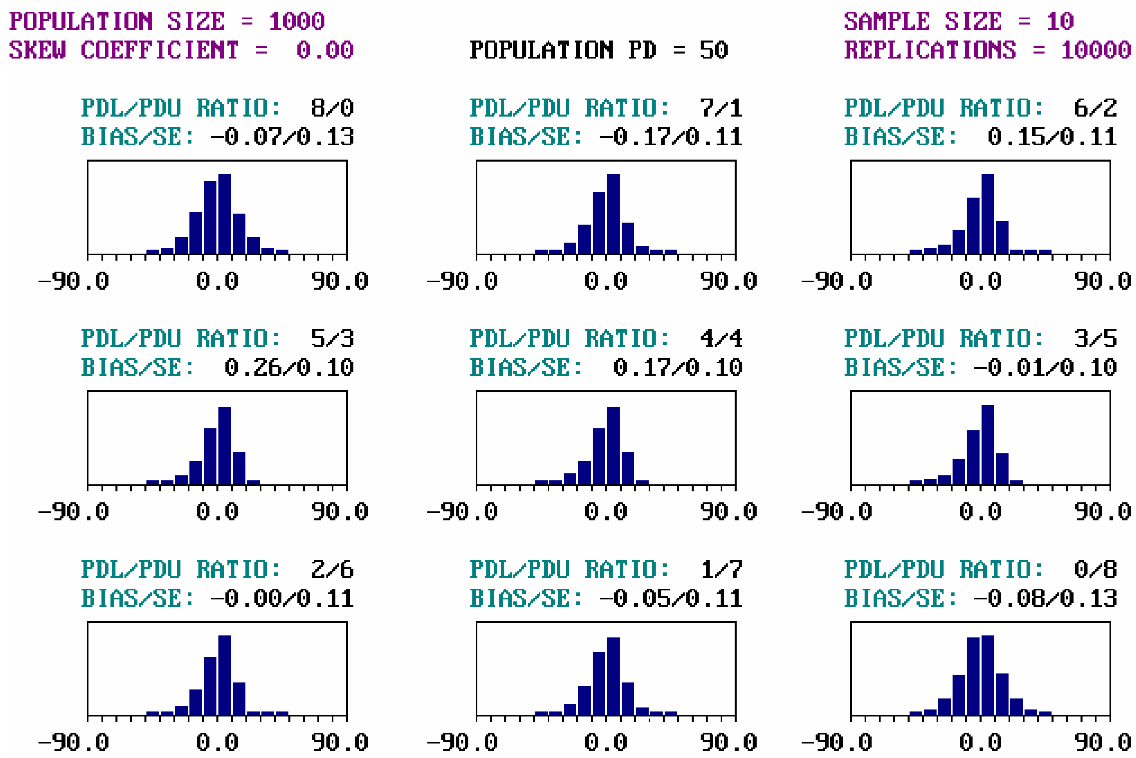


Figure 123. Sample output screen for a population with PD = 50, skewness coefficient = 0.00, and sample size = 10.

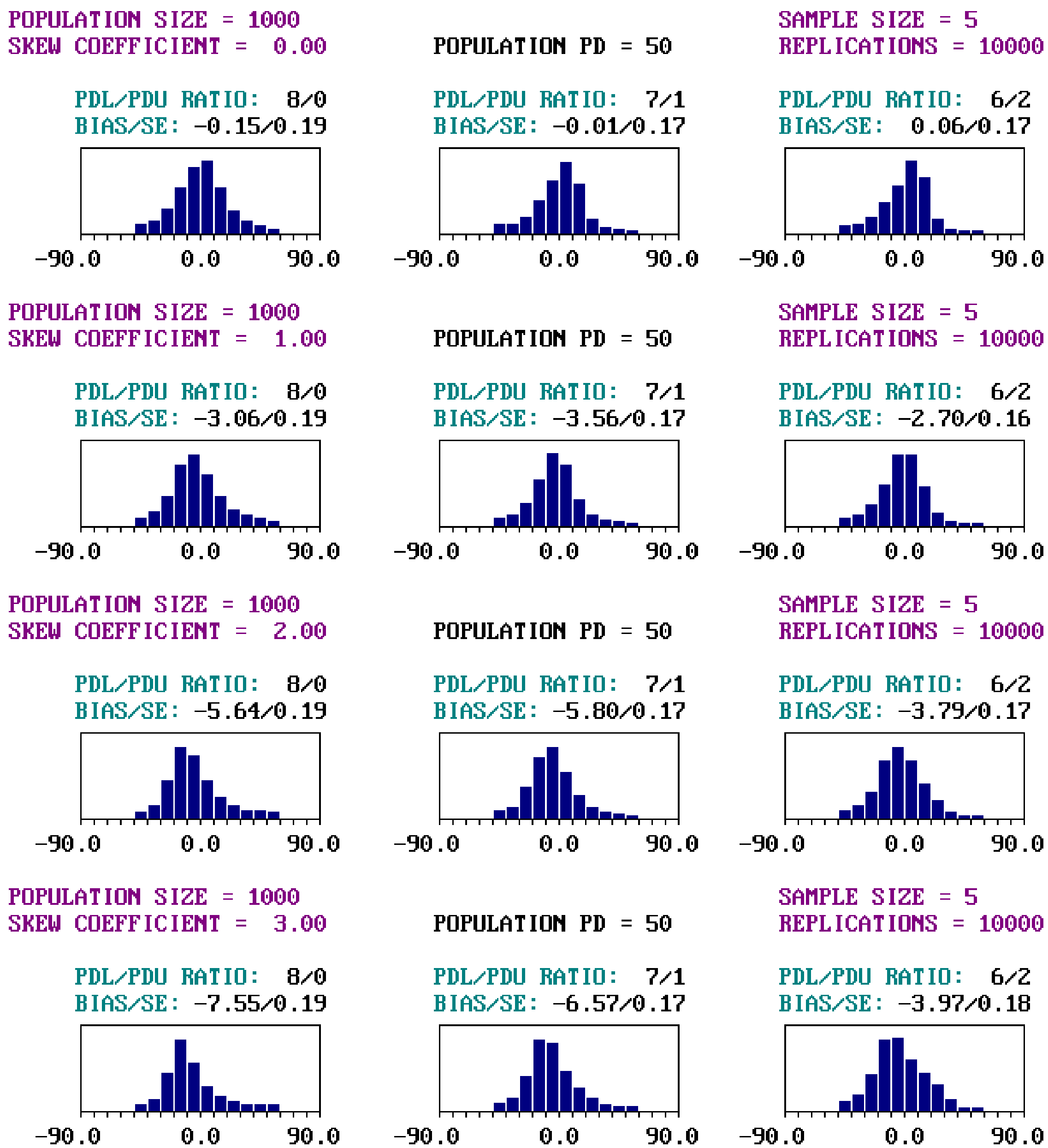


Figure 124. Portions of output screens for PD = 50, sample size = 5, and skewness coefficients = 0.00, 1.00, 2.00, and 3.00.

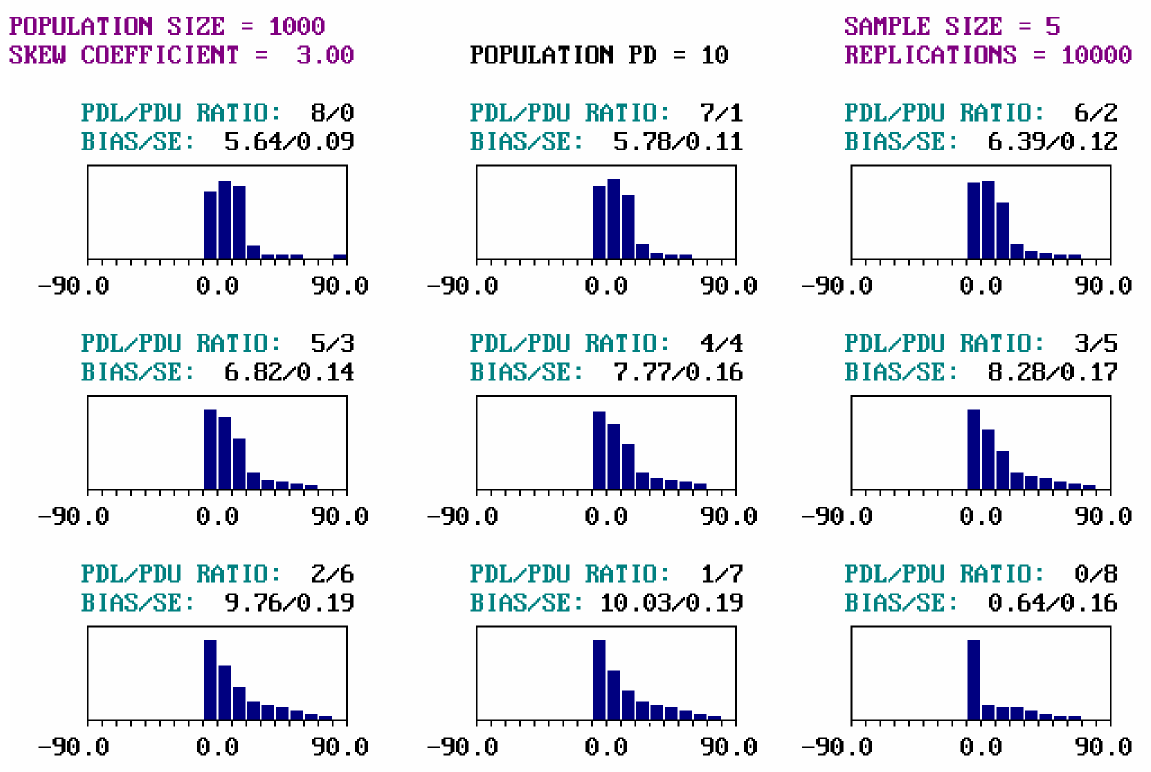


Figure 125. Sample output screen for a population with PD = 10, skewness coefficient = 3.00, and sample size = 5.

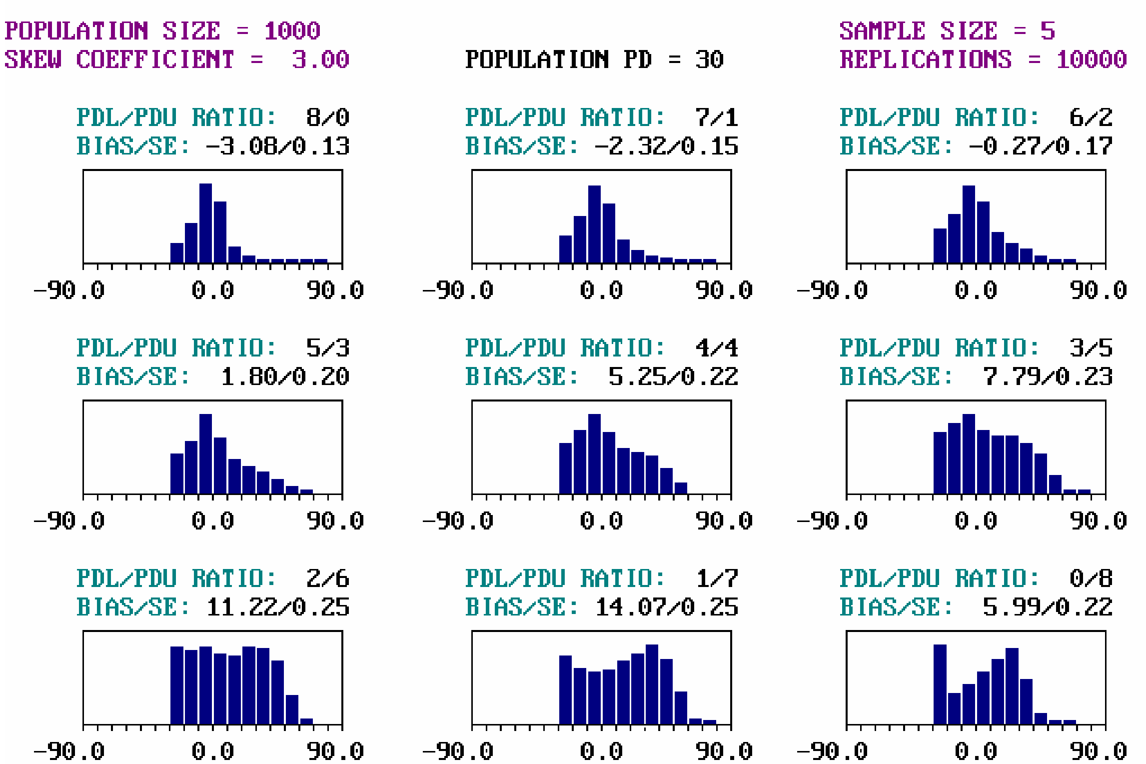


Figure 126. Sample output screen for a population with PD = 30, skewness coefficient = 3.00, and sample size = 5.

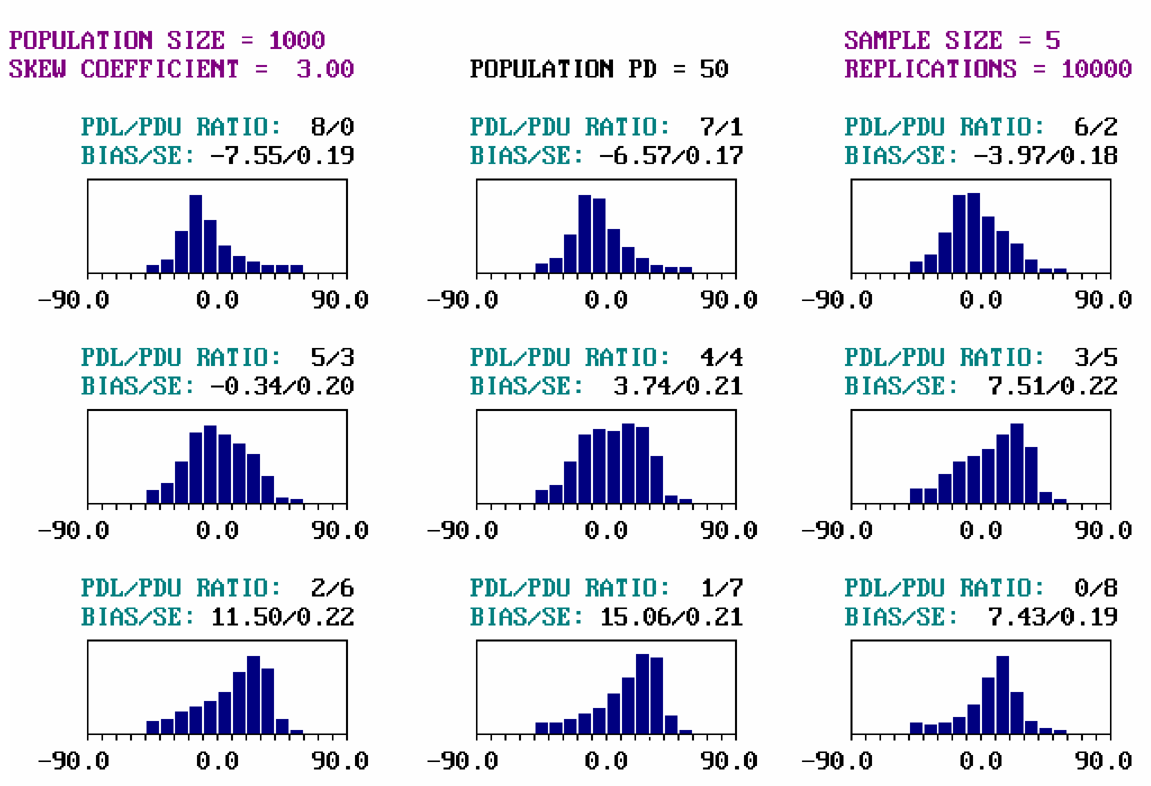


Figure 127. Sample output screen for a population with PD = 50, skewness coefficient = 3.00, and sample size = 5.

APPENDIX H: DISTRIBUTION OF SAMPLE PWL ESTIMATES

GENERATING SAMPLES

The plots in appendix H illustrate the distributions of sample PWL estimates for a number of different combinations of actual population PWL values. All of the plots are based on the same simulation results that came from simulating 1000 samples (sample size = 4) from a normal distribution with a mean of 0.00 and a standard deviation of 1.00. The sample means and standard deviations were then calculated for each of the 1000 samples. These values were then used to calculate sample PWL values for various actual population PWL values.

Figure 128 shows the distribution of the sample means for the 1000 samples. Statistical theory says that this should be a normal distribution with a mean equal to that of the population that was sampled (i.e., 0.00). Furthermore, the theory also says that the standard deviation of these 1000 sample means should be the population standard deviation (i.e., 1.00) divided by the square root of the sample size (i.e., $\sqrt{4} = 2$). This means that the standard deviation of the sample means should be 0.50.

The average for the 1000 simulated sample means was +0.01, while their standard deviation was 0.484. These values are close to the theoretical ones and they are well within the range that might be expected from a simulation study. Figure 128 shows distribution histograms for the sample means and the sample standard deviations. As expected, the distribution of the sample means is approximately normal and is centered around the population mean of 0.00. Also, as expected, the distribution of the sample standard deviations is obviously skewed to the right. The average of the 1000 sample standard deviations is 0.941, which is close to the population value of 1.00.

DISTRIBUTION OF SAMPLE PWL VALUES

The mean and standard deviation for each sample were used to estimate the PWL for populations with actual PWL values = 90, 80, 70, 60, and 50. For each population, both one-sided specifications and two-sided specifications were evaluated. For the two-sided specifications, the population mean was set in the middle of the specification range so that there were equal PD values for the lower and upper specification limits.

Figures 129 through 133 show the sample PWL distributions for both one-sided and two-sided populations with 90, 80, 70, 60, and 50 PWL, respectively. A trend regarding the shape of the distributions is apparent. For the population with its mean farthest from the specification limit(s) (i.e., the 90-PWL population), the distribution of the sample PWL values is skewed to the left. This is seen by the long tail to the left and the fact that about 60 percent of the distribution is to the right of the population PWL of 90. As the population mean moves closer to the specification limit(s) (i.e., moves from 80 to 70 to 60 to 50), skewness decreases until the distribution becomes symmetrical.

The reason that the sample PWL distributions are skewed stems from the fact that the distribution of the sample standard deviations is skewed. This skewness has less impact as the mean of the population approaches the specification limit. And, for a one-side specification to have 50 PWL, the population mean must be centered on the specification limit. Any population that is centered on the one specification limit will have 50 PWL regardless of the value of its standard deviation. So, in this instance, the distribution of the sample PWL values will be symmetrical about the population mean. For any sample mean that is below the population mean, the estimated PWL will be less than 50, while any sample mean above the population mean will have an estimated PWL greater than 50. This is shown in figures 128 and 133 where both have 48.9 percent below and 51.1 percent above their population values.

Figure 134 shows diagrams of the populations that were sampled. These populations can be compared with the distributions of the sample PWL values shown in figures 129 through 133. For the one-sided specifications, the distribution whose mean is farthest from the specification limit (i.e., 90 PWL) has more PWL values above the true value than below it. As the distance between the population mean and the specification limit decreases, the distributions become symmetrical about the true population value.

For the two-sided specifications, once again, the population whose mean is farthest from the specification limits (i.e., 90 PWL) has more PWL values above the true value than below it. However, as the population means come closer to the specification limits (i.e., 60 PWL and 50 PWL), the distributions have more values below the true value than above it.

The two situations that were simulated (one-sided specifications and two-sided specifications, with the populations centered between the two limits) are the extremes with respect to how the PD for a population with a given PWL value can be distributed outside the two limits. There are obviously many other ways in which these PWL values could be located with respect to the specification limits. From these examples, it appears that the shape of the distribution of the sample PWL values will vary depending on how the PWL range of the population is located inside the limit or limits.

However, regardless of the shape of the distribution of the sample PWL values that were simulated, the average of the 1000 sample PWL values all differed from the population PWL value by no more than about 0.5 percent. Although the EP values were all very close for each of the distributions, the different shapes indicate that there may be small differences in the percentage of time that a population PWL value is estimated to be too high and too low, depending on the PWL of the population and where the population is centered with respect to the specification limit or limits.

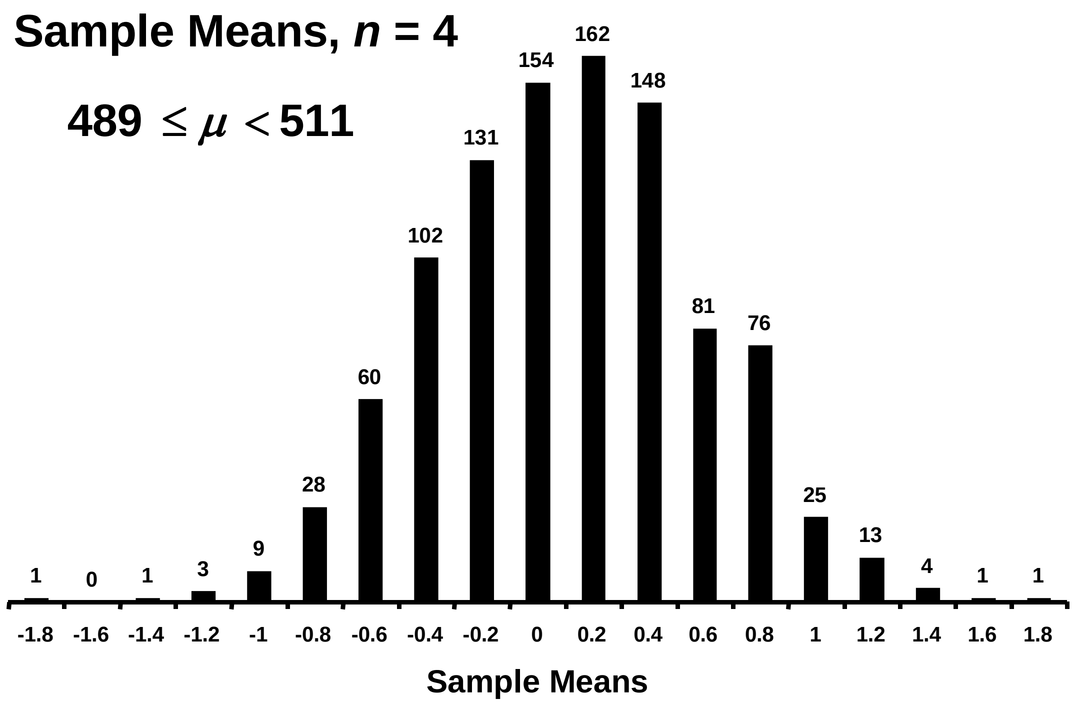


Figure 128a. Distribution of sample means for 1000 samples from a normal population with μ = 0.00, σ = 1.00.

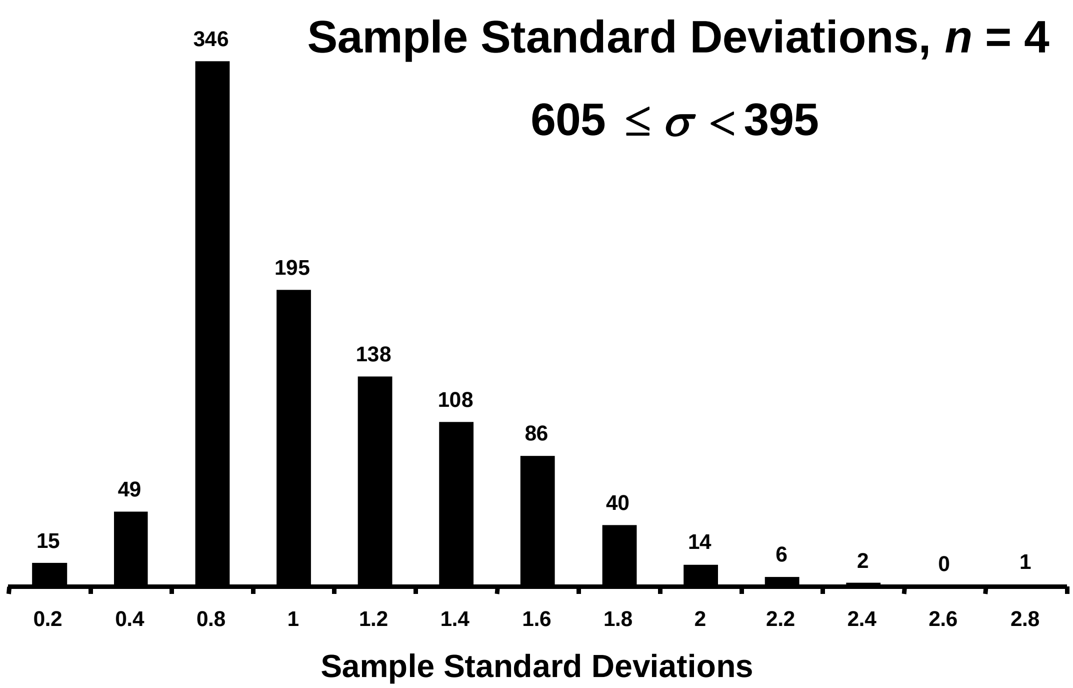


Figure 128b. Distribution standard deviations for 1000 samples from a normal population with μ = 0.00, σ = 1.00.

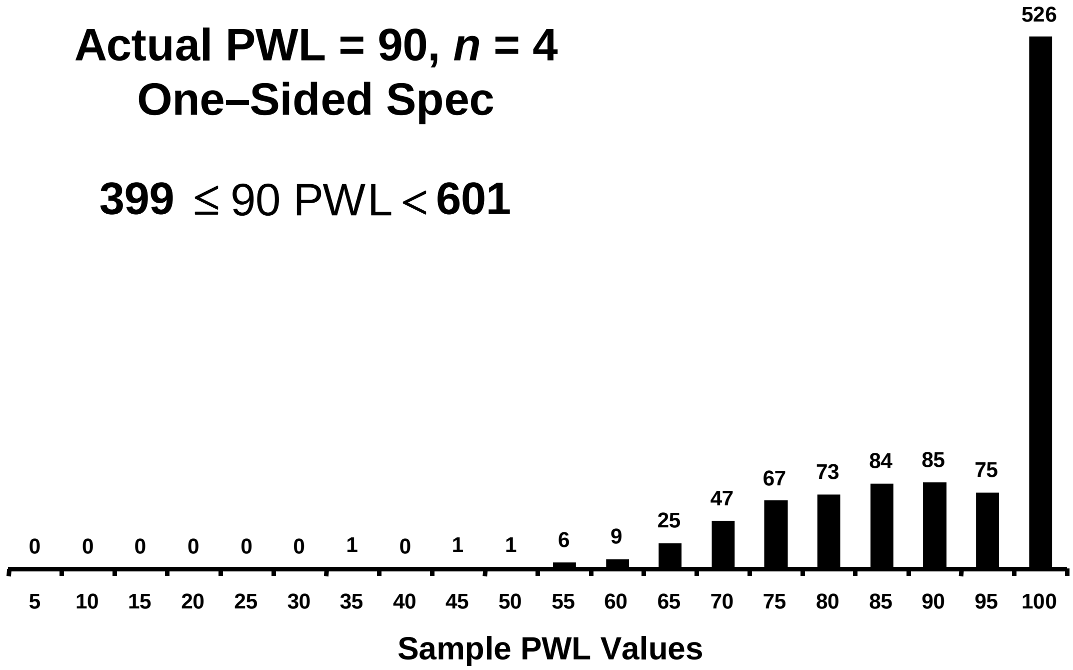


Figure 129a. Distribution of sample PWL values for 1000 samples from a normal population with PWL = 90, one-sided specifications.

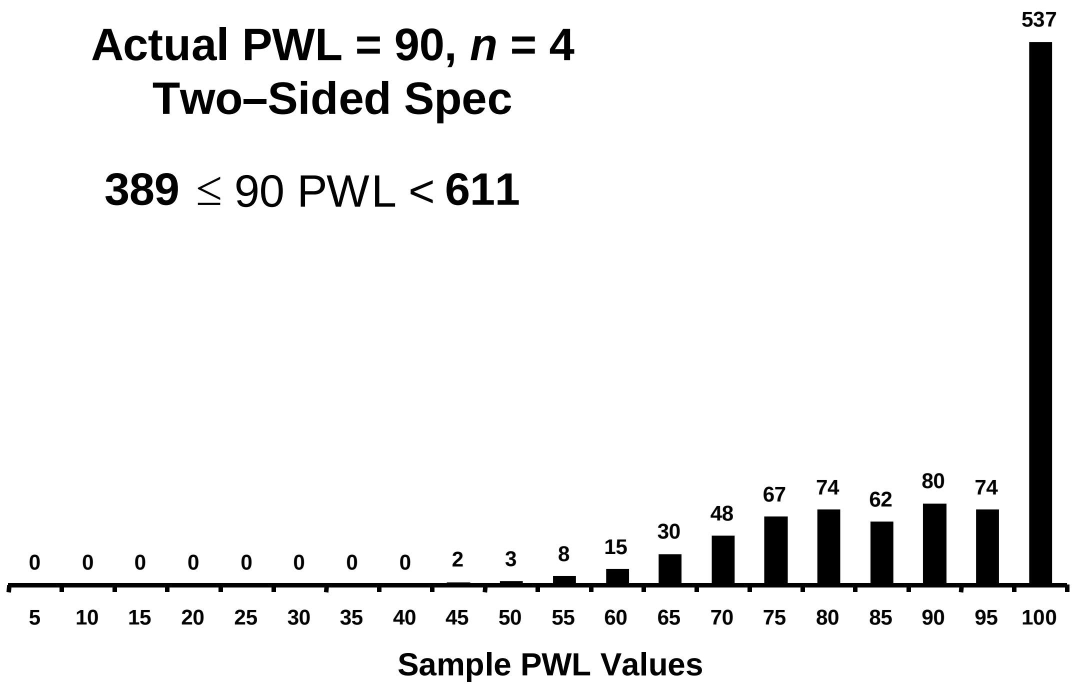


Figure 129b. Distribution of sample PWL values for 1000 samples from a normal population with PWL = 90, two-sided specifications.

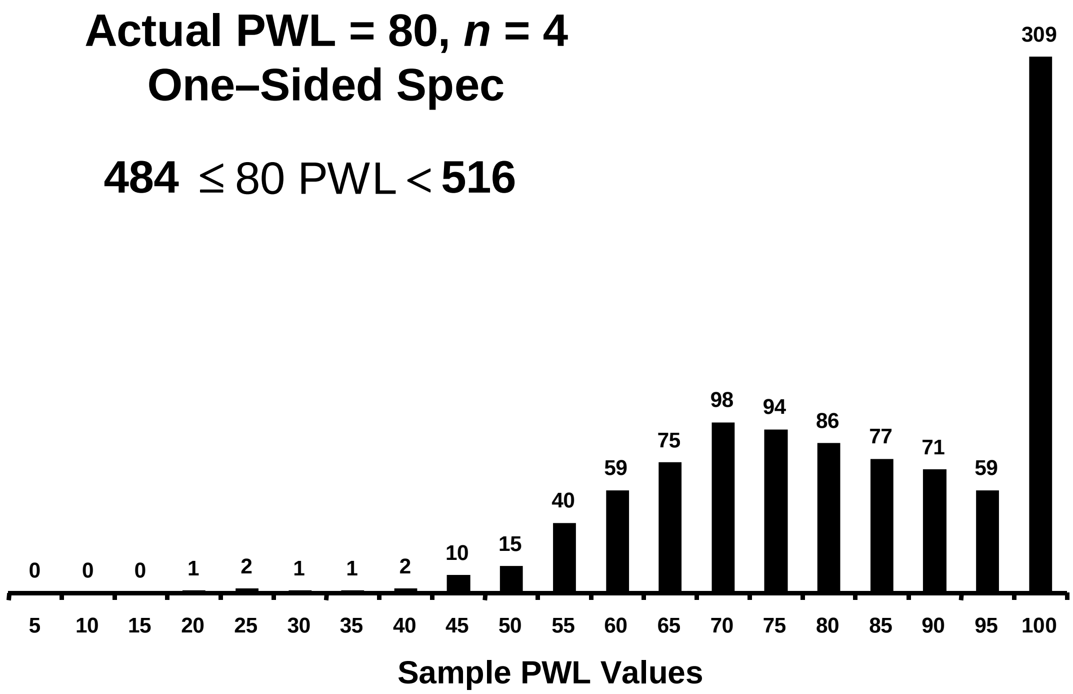


Figure 130a. Distribution of sample PWL values for 1000 samples from a normal population with PWL = 90, one-sided specifications.

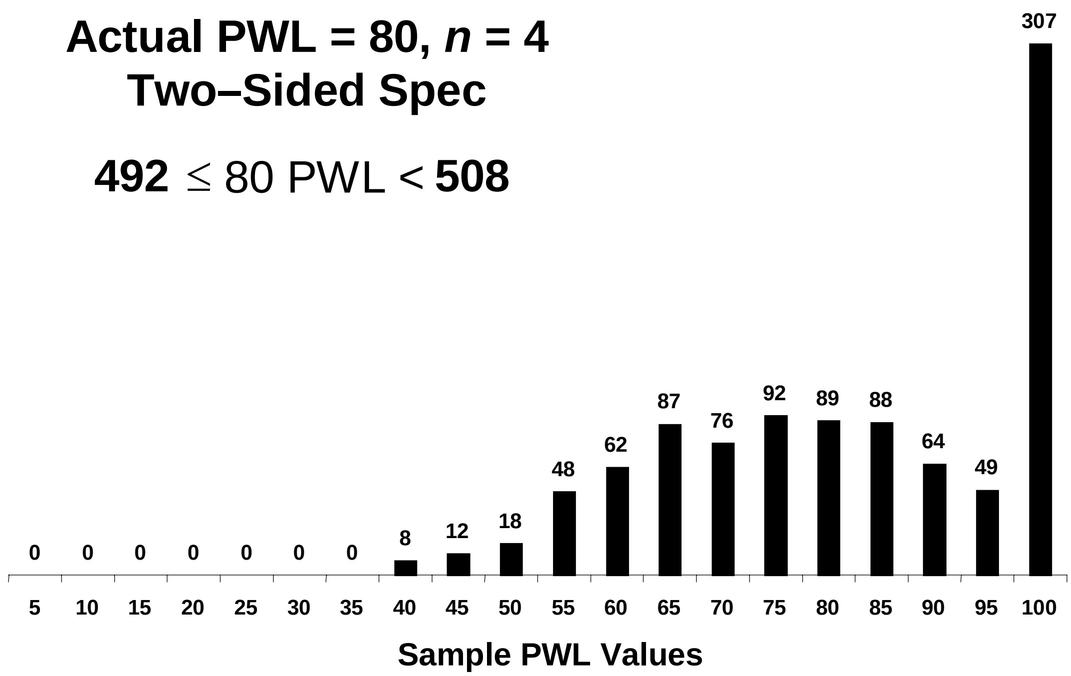


Figure 130b. Distribution of sample PWL values for 1000 samples from a normal population with PWL = 80, two-sided specifications.

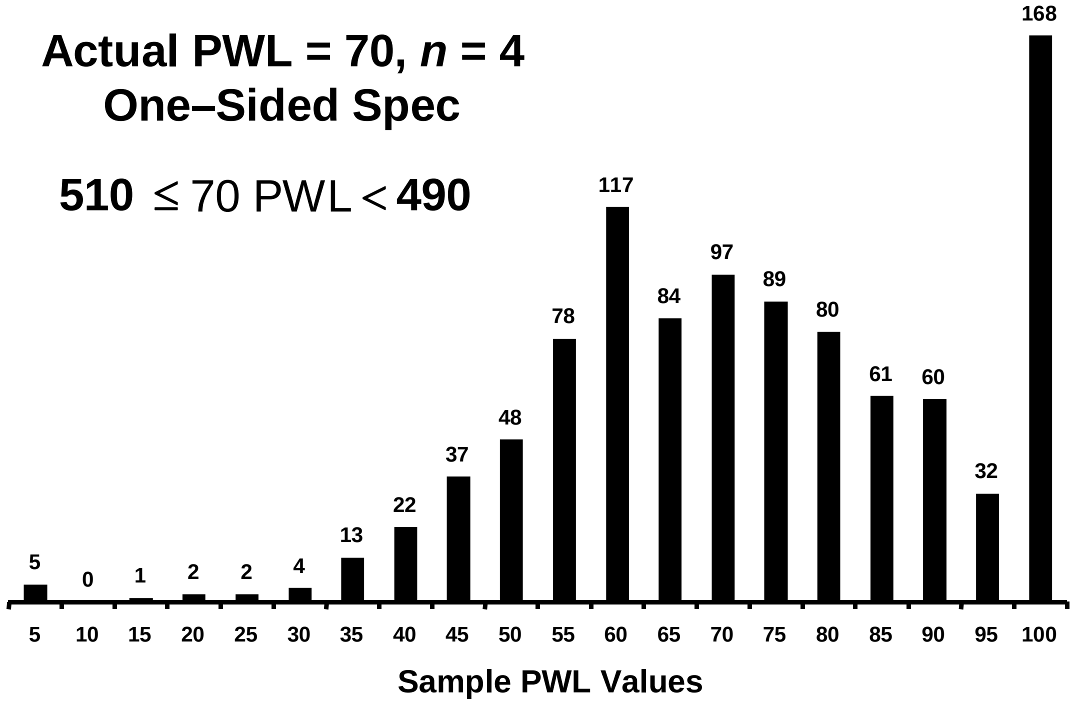


Figure 131a. Distribution of sample PWL values for 1000 sample from a normal population with PWL = 80, one-sided specifications.

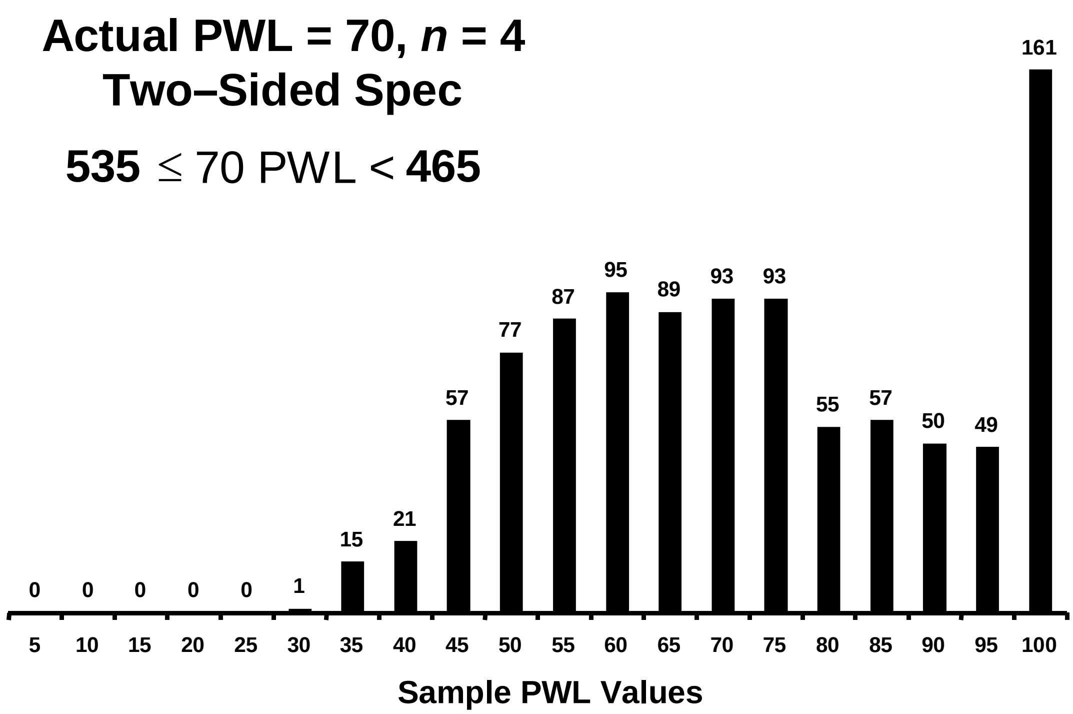


Figure 131b. Distribution of sample PWL values for 1000 samples from a normal population with PWL = 80, two-sided specifications.

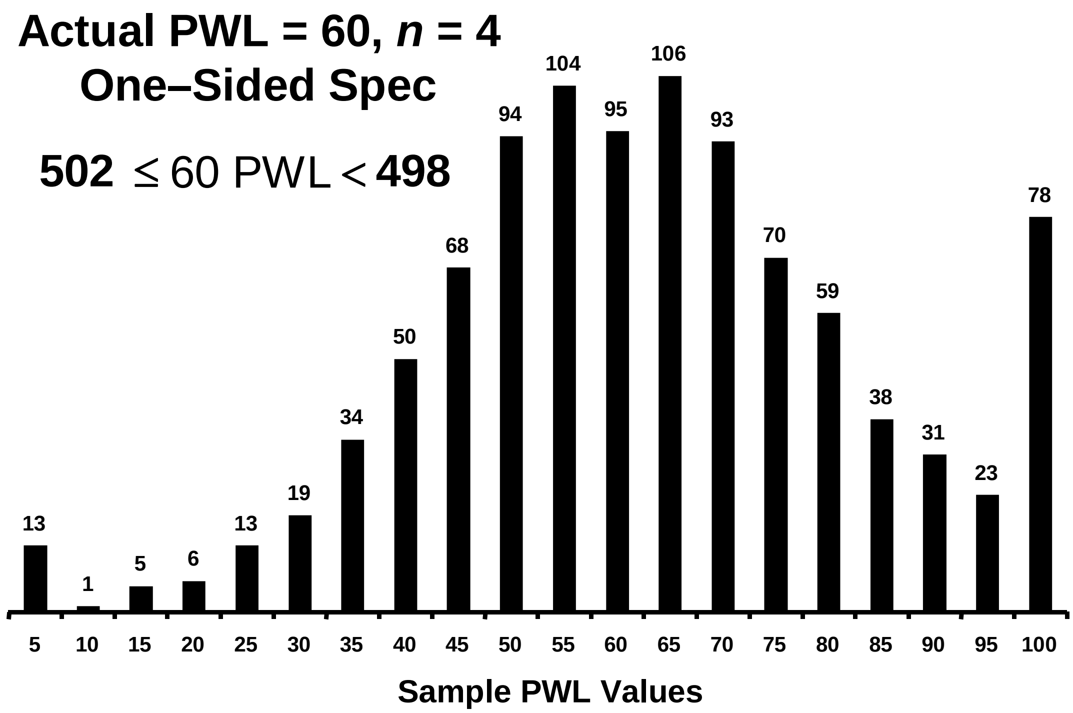


Figure 132a. Distribution of sample PWL values for 1000 samples from a normal population with PWL = 60, one-sided specifications.

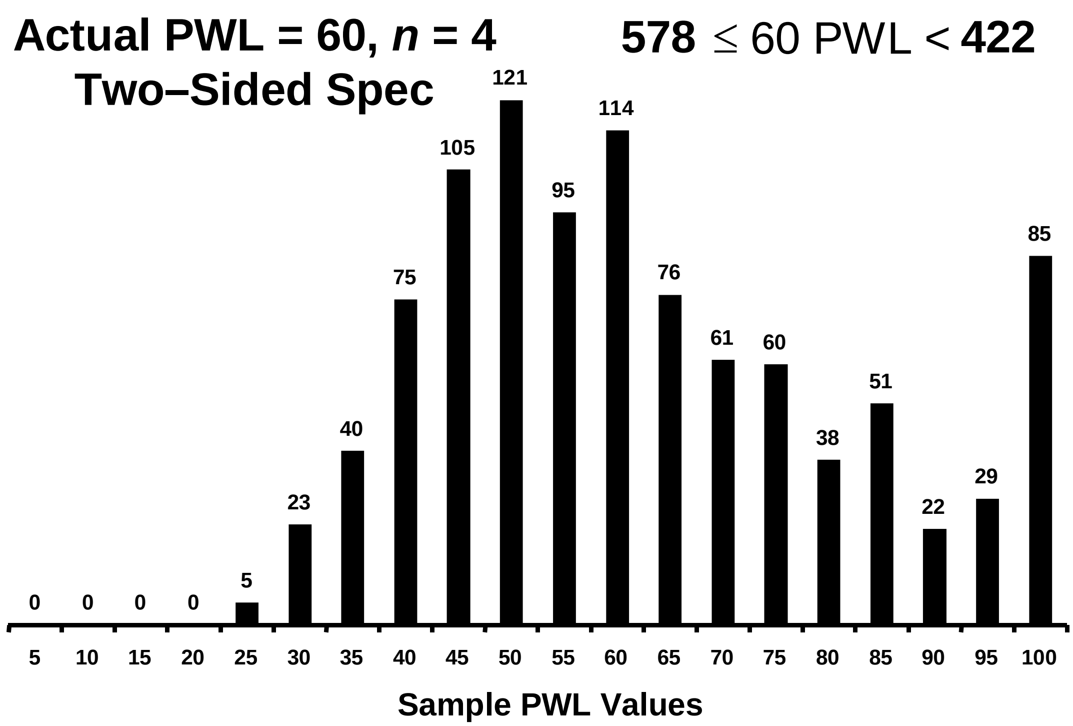


Figure 132b. Distribution of sample PWL values for 1000 samples from a normal population with PWL = 60, two-sided specifications.

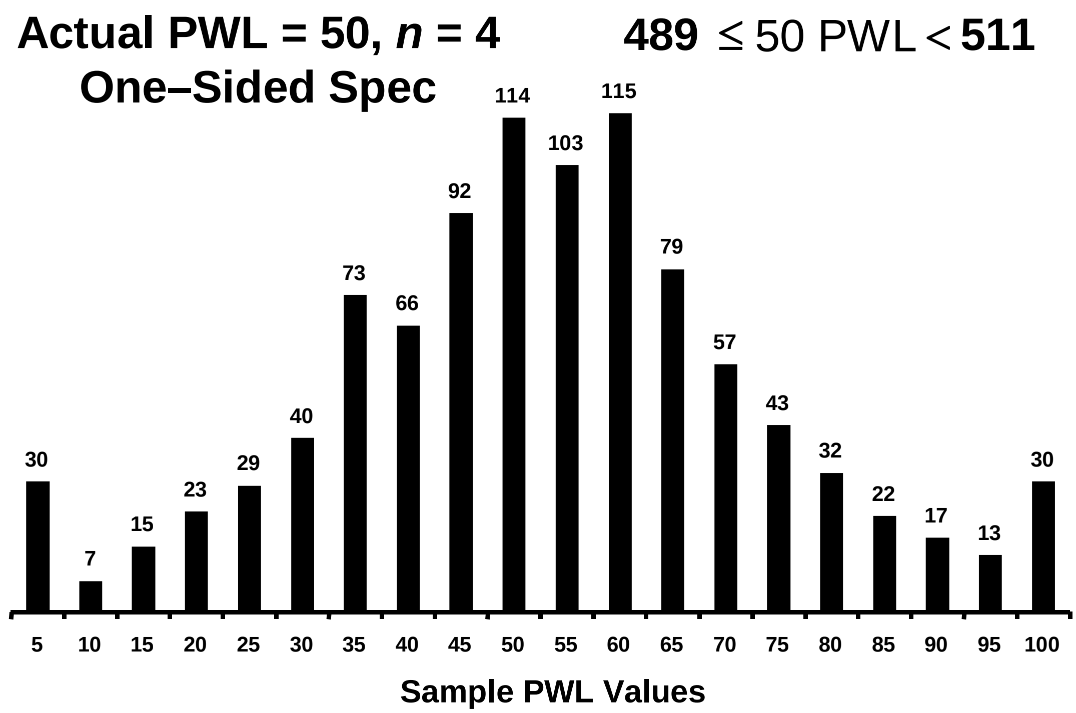


Figure 133a. Distribution of sample PWL values for 1000 samples from a normal population with PWL = 50, one-sided specifications.

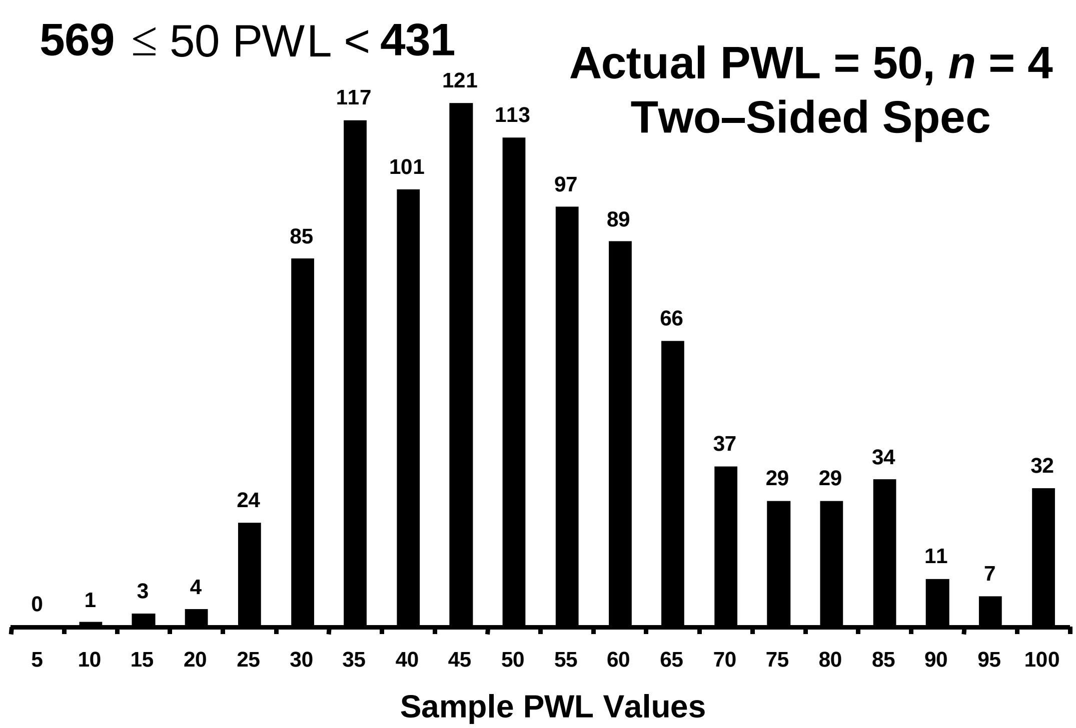


Figure 133b. Distribution of sample PWL values for 1000 samples from a normal population with PWL = 50, two-sided specifications.

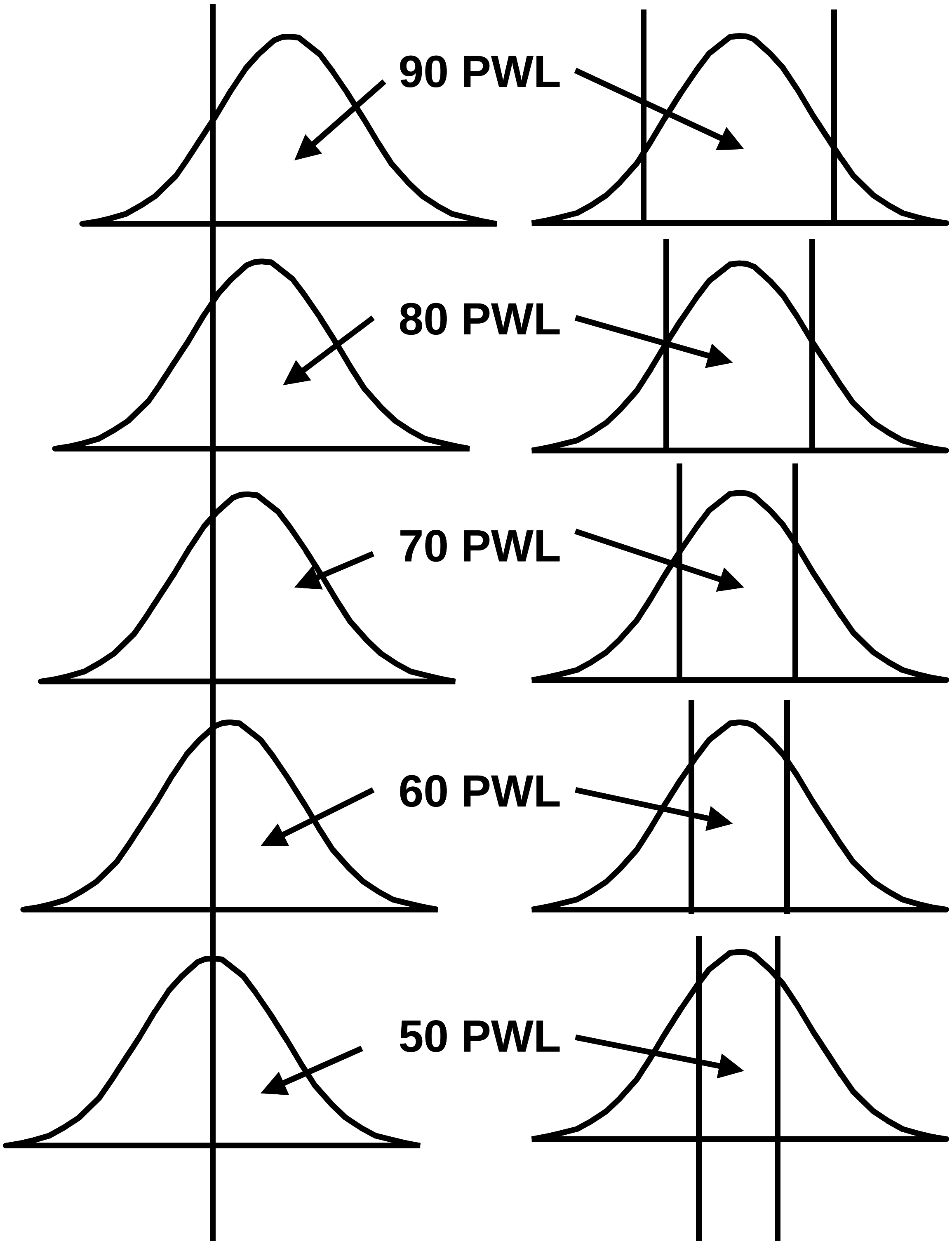


Figure 134. Illustration of the populations for which the distributions of sample PWL values are shown in figures 129 through 133.

REFERENCES

1. Burati, J.L., et al. *Optimal Procedures for Quality Assurance Specifications*, Publication No. FHWA-RD-02-095, Federal Highway Administration, Washington, DC, April 2003.

2. *AASHTO Implementation Manual for Quality Assurance*, AASHTO Quality Construction Task Force, American Association of State Highway and Transportation Officials, February 1996.

3. *AASHTO Quality Assurance Guide Specification*, AASHTO Quality Construction Task Force, American Association of State Highway and Transportation Officials, February 1996.

4. "Glossary of Highway Quality Assurance Terms," *Transportation Research Circular No. E-C037*, Transportation Research Board, Washington, DC, April 2002.

Recycled
Recyclable

HRDI-11/10-04(800)E